

RESULT 2
ENTRY B37314 #type complete
TITLE regulatory protein lcrV - Yersinia pseudotuberculosis
ORGANISM #formal_name Yersinia pseudotuberculosis
DATE 30-Jan-1993 #sequence_revision 30-Jan-1993 #text_change 17-Mar-1999

ACCESSIONS
REFERENCE B37314
#authors Bergman, T.; Hakansson, S.; Forsberg, A.; Norlander, L.; Macellaro, A.; Baeckman, A.; Boellin, I.; Wolf-Watz, H.
#title J. Bacteriol. (1991) 173:1607-1616
Analysis of the V antigen lcrGV-yopBD operon of Yersinia pseudotuberculosis: evidence for a regulatory role of lcrH and lcrV.

#cross-references MUID:91154114
#accession B37314
#status Preliminary
#molecule_type DNA
#residues 1-326 #label BER
#cross-references GB:M57893; NID:q155456; PID:q155458
SUMMARY #length 326 #molecular-weight 37336 #checksum 4689

Query Match 61.8%; Score 2192; DB 2; Length 326;
Best Local Similarity 96.6%; Pred. No. 0.00e+00;
Matches 315; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

DB 1 MIRAEONPOHFIEDEKRVFOLTGHSVLELVOLVKDNIIDSIKYPDKSEVFA 60
OY 196 MIRAEONPOHFIEDEKRVFOLTGHSVLELVOLVKDNIIDSIKYPDKSEVFA 255
DB 61 NRVITDDELLKRIIAYFLPEDAIKGGHYDQNLONGIKRVKEFLESSPNTOMELRAEFA 120
OY 256 NRVITDDELLKRIIAYFLPEDTIKGGHYDQNLONGIKRVKEFLESSPNTOMELRAEFA 315
DB 121 VHFSLTADRIDDDILKIVYDSMNHGDAKRSKRLREELAEIETAEIKYVYQAEINKHLS 180
OY 316 VHFSLTADRIDDDILKIVYDSMNHGDAKRSKRLREELAEIETAEIKYVYQAEINKHLS 375
DB 181 GGTINIHOKSINLMDKNLYGYDEDEFKASAEYKILEKMPOTTIOGGEKKIVSKNLT 240
OY 376 GGTINIHOKSINLMDKNLYGYDEDEFKASAEYKILEKMPOTTIOGGEKKIVSKNLT 435
DB 241 ESEKRTGAGNLKDSYSYKNKNDELSPATTCSDKSRLPLDVSKTQLSDITSRENS 300
OY 436 GSEKRTGAGNLKDSYSYKNKNDELSPATTCSDKSRLPLDVSKTQLSDITSRENS 495
DB 301 AIEALNRFIOKIDSVQRLDDTSGK 326
OY 496 AIEALNRFIOKIDSVQRLDDTSGK 521

RESULT 3
ENTRY S13008 #type complete
TITLE capsular antigen F1 precursor - Yersinia pestis
ORGANISM #formal_name Yersinia pestis
DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 13-Sep-1998
ACCESSIONS S13008; S23725
REFERENCE S13008
#authors Gal'gov, E.E.; Smirnov, O.Y.; Karlishhev, A.V.; Volkovoy, K.I.; Denesvuk, A.I.; Nazimov, I.V.; Rbltsov, K.S.; Abramov, V.M.; Dalvadganz, S.M.; Zav'yalov, V.P.
#title FEBS Lett. (1990) 277:230-232
Nucleotide sequence of the Yersinia pestis gene encoding F1 antigen and the primary structure of the protein. Putative T and B cell epitopes.

#cross-references MUID:91099503
#accession S13008
#molecule_type DNA
#residues 1-170 #label GAL
#cross-references EMBL:X61996; NID:g48620; PID:g48621

GENETICS
#gene cafi
FEATURE 1-21 #domain signal sequence #status predicted #label SIGY
22-170 #product capsular antigen F1 #status predicted #label MAT

SUMMARY #length 170 #molecular-weight 17666 #checksum 5455

Query Match 31.0%; Score 1100; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.52e-139;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MKKISSVIAIAFGTIAANAAADLTASTATATLVEPRRIITLYEGAPITIMDNGIDT 60
OY 24 MKKISSVIAIAFGTIAANAAADLTASTATATLVEPRRIITLYEGAPITIMDNGIDT 83
DB 61 ELLVGTFLGKYKGTSTSVNFETDAADPMYLFPTTSODGNHOTTGVICKDSRDPIS 120
OY 84 ELLVGTFLGKYKGTSTSVNFETDAADPMYLFPTTSODGNHOTTGVICKDSRDPIS 143
DB 121 PKVNGENLVGDVVLATSGODFFVRSISGKGLAKGYTAVTYVSNQ 170
OY 144 PKVNGENLVGDVVLATSGODFFVRSISGKGLAKGYTAVTYVSNQ 193

RESULT 4
ENTRY I58315 #type complete
TITLE WT1 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Aug-1997

ACCESSIONS I58315
REFERENCE I58315
#authors Hamilton, T.B.; Barilla, K.C.; Romanuk, P.J.
#title Nucleic Acids Res. (1995) 23:277-284
High affinity binding sites for the Wt1s' tumour suppressor protein WT1.

#cross-references MUID:95166649
#accession I58315
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-168 #label RES
#cross-references GB:S75264; NID:g896246; PID:g896247

GENETICS
#gene GDB:WT1
#cross-references GDB:120496; OMIM:194070
#map_position 11p13-11p13
SUMMARY #length 168 #molecular-weight 20165 #checksum 6457

Query Match 4.7%; Score 168; DB 2; Length 168;
Best Local Similarity 66.7%; Pred. No. 9.77e-06;
Matches 20; Conservative 7; Mismatches 1; Indels 2; Gaps 1;

DB 1 MGNHHHHHHHSGHICG--RHMRPVCV 28
OY 1 MGNHHHHHHHSGHIDDDKMKKISSV 30

RESULT 5
ENTRY S57382 #type complete
TITLE hypothetical protein YOL087c - yeast (Saccharomyces cerevisiae)
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 12-Dec-1997
ACCESSIONS S57382; S56781; S50418
REFERENCE S57374
#authors Zurewtein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
#title Yeast (1995) 11:975-986
A 29.425 kb segment on the left arm of yeast chromosome XV contains more than twice as many unknown as known open

[illegible]

QY 93 GCYKGTSTSVNFTDAAGDPMLTFTSDGNNHOFTTKVIGKDSRDPDISPKVNGENLV 152

Db 81 ID 82

QY 153 GD 154

RESULT 8

ENTRY S33821 #type complete

TITLE median body protein - Giardia lamblia

ORGANISM #formal_name Giardia lamblia

DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997

ACCESSIONS S33821

REFERENCE S33821

#authors Marshall, J.; Holberton, D.V.

#journal J. Mol. Biol. (1993) 231:521-530

#title Sequence and structure of a new coiled coil protein from a microtubule bundle in Giardia.

#cross-references MUID:93287123

#accession S33821

#status Preliminary

#molecule_type mRNA

#residues 1-857 #label MAR

SUMMARY #cross-references EMBL:X64517; NID:g312670; PID:g312671

#length 857 #molecular_weight 100583 #checksum 6805

Query Match 3.5%; Score 124; DB 2; Length 857;

Best Local Similarity 20.1%; Pred. No. 4,618-01;

Matches 43; Conservative 71; Mismatches 82; Indels 18; Gaps 17;

Db 582 ERVVEERIQEKENNLKQRTLEEQQRTATVRETEMSALREKANELDGYNRRQAREH 641

QY 205 QHFTED-LEKRYVEQLTGSSVLEELVOLVKDKNIDISIKYDPKDESEFA-NRVID- 261

Db 642 EIMMLDALKSDKLRONRWAMEL-TELREKVOLEKLOEKARARVEMLELRKAMD 700

QY 262 DIELK-KILAYF-LPEPTILKGGHYDNOLONGIKRV-K-EFLESSPMTQ-WELFA-FMA 315

Db 701 VDTLVEEQKQRLAEIKVNNYDQADKARLOEQLEKMSDKLIFEMIMD-NRRL 759

QY 316 VMHFSLTADRIDDDI--LKIVYDSNMHHGDARSKRELAELTALKTIVYQAEINKHL 373

Db 760 KLOVK-ELDLKTAN-MEK-LYEEYKLEDDQKAT 790

QY 374 SSSGTHIHDKSINIMDKNLVG-YTD-EEIFKAS 405

RESULT 9

ENTRY S15999 #type complete

TITLE fatty-acyl-CoA synthase (EC 2.3.1.86) beta chain - yeast

ORGANISM (Yarrowia lipolytica)

DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998

ACCESSIONS S15999

REFERENCE S15999

#authors Koeltig, H.; Rotner, G.; Beck, K.F.; Schweizer, M.; Schweizer, E.

#journal Mol. Gen. Genet. (1991) 226:310-314

#title The pentafunctional FAS1 genes of Saccharomyces cerevisiae and Yarrowia lipolytica are co-linear and considerably longer than previously estimated.

#cross-references MUID:91238709

#accession S15999

GENETICS #molecule_type DNA

#residues 1-2076 #label MOL

KEYWORDS #cross-references EMBL:X59690; NID:g297854; PID:g297855

CLASSIFICATION FAS1

SUMMARY #superfamily yeast fatty-acyl-CoA synthase beta chain

#keywords acyltransferase

#length 2076 #molecular_weight 230223 #checksum 5483

Query Match 3.5%; Score 123; DB 2; Length 2076;

Best Local Similarity 31.3%; Pred. No. 5,778-01;

Matches 25; Conservative 22; Mismatches 28; Indels 5; Gaps 5;

Db 1268 TDRNTRIKFEFYKLMFGODSK-FEIDTDIT-ELIGSDVYITSGKALDFVHANGKEAF 1325

QY 119 TSQGNHNOFTTKV-IGKDSRDPDISPKVNGENLVGDVVLATGSDQDFVFRISIGSKGKL 177

Db 1326 V-GRSTAGFY-FAPMDFAI 1343

QY 178 AAGKTYDAVTVYSNOEFMI 197

RESULT 10

ENTRY S48450 #type complete

TITLE hypothetical protein YIL007c - yeast (Saccharomyces cerevisiae)

ORGANISM #formal_name Saccharomyces cerevisiae

DATE 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 12-Dec-1997

ACCESSIONS S48450

REFERENCE S48442

#authors Rowley, N.

#submission submitted to the EMBL Data Library, August 1994

#accession S48450

#molecule_type DNA

#residues 1-220 #label ROW

SUMMARY #cross-references GB:Z47047; EMBL:Z38113; NID:g603997; PID:g763339; MIPS:YIL007c

GENETICS #map_position 9L

Query Match 3.4%; Score 120; DB 2; Length 220;

Best Local Similarity 22.8%; Pred. No. 1,128+01;

Matches 34; Conservative 39; Mismatches 69; Indels 7; Gaps 7;

Db 38 KTDIETOLEAV-FSVLEQOC-IGMDSALVTPDGY-PRSDVDVQLQVTMIRKNVNLKNDLN 94

QY 213 KVRVE-QLTGSSVLEELVOLVKDKNIDISIKYDPKDESEFANRYITDDIELKKILA 271

Db 95 HILORSHVLLNQHFDDNMVKSNDARRN-NDQAIOYTIPTAFISEVVPSPSDKADIKV 153

QY 272 YFLPEDYILKGGHYDNOLONGIKRVEFLESSPMTQWELR-AFMAVHFSLTADRIDDDI 330

Db 154 DDKLISTGNYHAHNSKLNQWVVMKNE 182

QY 331 LKVIYDSNMHHGDARSKLRE-ELAEITAE 358

RESULT 11

ENTRY H64591 #type complete

TITLE signal peptidase I Helicobacter pylori (strain 26955)

ORGANISM #formal_name Helicobacter pylori

DATE 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 10-Oct-1997

ACCESSIONS H64591

REFERENCE A64520

#authors Tomb, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klein, H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.; McKenney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.M.; Venter, J.C.

#journal Nature (1997) 388:539-547

#title The complete genome sequence of the gastric pathogen Helicobacter pylori.

```

#cross-references MUID:97394467
#accession H64591
#status preliminary; nucleic acid sequence not shown;
translational not shown
##molecule_type DNA
#residues 1-290 ##label TOM
##cross-references GB:AE00571; GB:AE00511; NID:g2313686; PID:g2313695,
TIGR:HP0576
SUMMARY #length 290 #molecular-weight 33710 #checksum 897

Query Match 3.4% Score 120; DB 1; Length 290;
Best Local Similarity 27.2%; Pred. No. 1.12e+00;
Matches 25; Conservative 28; Mismatches 32; Indels 7; Gaps 7;

Db 111 ICGDEVLT-NEGFTLHPESDPTDKNTAKHYNNATKEFMGKIFVLANYNKNEHPGHHQ 169
:| | | | | :| | | | | :| | | | | :| | | | |
Oy 152 VGDDVVLATGSDPFVNASISGKGLAKG-YTDATVTVVSNQEFNIRAY-EQNPQ-HFI 208
:| | | | | :| | | | | :| | | | | :| | | | |
Db 170 KDNETHLMEOATQGAENANISMQIOMEKG 201
:| | | | | :| | | | | :| | | | | :| | | | |
Oy 209 EDLEKVR-VEQLTGHGSSV-LE-ELVQLVKDK 237

RESULT 12
ENTRY KIBYRB #type complete
TITLE ribokinase (EC 2.7.1.15) - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein YCR036w; protein YCR523
ORGANISM #format_name Saccharomyces cerevisiae
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
12-Dec-1997
ACCESSIONS S12918; S19448
REFERENCE S12916
#authors Thierry, A.; Fairhead, C.; Dujon, B.
#journal Yeast (1990) 6:521-534
#title The complete sequence of the 8.2 kb segment left of MAT on
chromosome III reveals five ORFs, including a gene for a
yeast ribokinase.
#cross-references MUID:91181345
#accession S12918
##molecule_type DNA
##residues 1-333 ##label THI
##cross-references EMBL:X56909; NID:g44489; PID:g44492
REFERENCE S19445
#authors Herbert, C.J.; Jia, Y.; Slonimski, P.P.
#submission submitted to the Protein Sequence Database, March 1992
#accession S19448
##molecule_type DNA
##residues 1-333 ##label D0J
##cross-references EMBL:X59720; NID:g1907116; PID:e264378; PID:g1907178,
MIPS:YCR036w

GENETICS
#gene SGD:RBK1
##cross-references SGD:S0000632; MIPS:YCR036w
CLASSIFICATION #map_position 3R
KEYWORDS #superfamily ribokinase
phosphotransferase
SUMMARY #length 333 #molecular-weight 36984 #checksum 5645

Query Match 3.4% Score 120; DB 1; Length 333;
Best Local Similarity 23.2%; Pred. No. 1.12e+00;
Matches 22; Conservative 25; Mismatches 47; Indels 1; Gaps 1;

Db 235 KLMNRKRGIVYVMTLSRGVLFCSHSEPEVQLPAIONYSYVDITGAGDTFLGGLYQLX 294
:| | | | | :| | | | | :| | | | | :| | | | |
Oy 176 KLAAGKTTDAVTVVSNQEFNIRAYEQNPQHTEDLEKRVEDLTGHGSSVLEELV-QLV 234
:| | | | | :| | | | | :| | | | | :| | | | |
Db 295 QGETLSMAIKFSTLASSLTQKGAESMPLYKDV 329
:| | | | | :| | | | | :| | | | | :| | | | |
Oy 235 KDKNIDISIKYDPRKDESEFANRVYTDIDLLAKTI 269

RESULT 13
ENTRY I58156 #type complete

```

```

TITLE      Brn-3.2 - mouse
ORGANISM   #formal_name Mus sp. #common_name mouse
DATE       26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change
          17-Oct-1997
ACCESSIONS 158156
REFERENCE   158156
AUTHORS     Turner, E.E.; Jenne, K.J.; Rosenfeld, M.G.
JOURNAL     Neuron (1994) 12:205-218
#title      Brn-3.2: a Brn-3-related transcription factor with
          distinctive central nervous system expression and
          regulation by retinoic acid.
#cross-references MUID:94121912
#accession 158156
#status     Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues   1-411 ##label RES
#cross-references GB:S68377; NID:9545068; PID:9545069
GENETICS
#gene       Brn-3.2
CLASSIFCATION #superfamily unassigned homeobox proteins; homeobox homology
KEYWORDS     POU domain homology
FEATURE       DNA binding; homeobox; nucleus; transcription regulation
259-329      #domain POU domain homology #label POU
348-404       #domain homeobox homology #label HOX
SUMMARY      #length 411 #molecular-weight 43172 #checksum 9634

Query Match      3.4%; Score 121; DB 2; Length 411;
Best Local Similarity 28.6%; Pred. No. 9,01e-01;
Matches 24; Conservative 22; Mismatches 35; Indels 3; Gaps 3;

Db 174 HHHHHHHHHHPQALGELLEHSPGLAGAMAGDGTIVSRPAHAPMTNPNHOA 233
||||||| : : : : : : : : : : : : : : : : : : : : : : : : :
OY 3 HHHHHHHHSSGRIHDDDKMKKRISSYIAL-ALFGIITATNAADLTASTTATATYPERA 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 62 RITLYKRGAPITIMD-NGNIDTE 84

RESULT 14
ENTRY   T01384 #type complete
TITLE   hypothetical protein T419.3 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
          cress
DATE     12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
          24-Mar-1999
ACCESSIONS T01384
REFERENCE   214314
AUTHORS     Parnell, L.D.; Gnoj, L.; de la Bastide, M.; Hameed, A.;
          Habermann, K.; Schutz, K.; Huang, E.; Gottesman, T.;
          Dedhia, N.N.; McCormick, W.R.
          Submitted to the EMBL Data Library, May 1998
          Genomic sequence of BAC T419 from Arabidopsis thaliana,
          chromosome IV, near 16.6 cm.
#accession T01384
#status     translated from GB/EMBL/DBJ
#molecule_type DNA
#residues   1-925 ##label PAR
#cross-references EMBL:AF069442; NID:g3242970; PID:g3924595
#experimental_source cultivar Columbia

GENETICS
#map_position 4
#introns      10/3; 54/3; 96/3; 148/3; 176/3; 219/3; 292/3; 336/3; 367/3;
          418/3; 515/1; 561/1; 584/2; 604/3; 761/3
          T419.3
SUMMARY      #length 925 #molecular-weight 102857 #checksum 8428

Query Match      3.4%; Score 120; DB 2; Length 925;
Best Local Similarity 17.4%; Pred. No. 1.12e+00;
Matches 16; Conservative 43; Mismatches 29; Indels 4; Gaps 4;

```

Qy	181	KYTLA-VIATVSNQEFMI-RAVEQNPOHFIE-DLEKXRVEDLIGHSSSVLEELVQLVKDK	237
Db	625	NEASEDCDVAKEG-VMTKGCVNEIDTVERL	655
Qy	238	NIDISIKYDPKKDSEVFANRKYITTDIELTKKI	269

ENTRY TITLE	RESULT	15
A56923 transcription factor shn - melanogaster)	#type complete	fruit fly (Drosophila

```

ALTERNATE_NAMES    schnurri
ORGANISM            #formal_name Drosophila melanogaster
DATE                11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change
                    17-Mar-1999

```

ACCESSIONS	A56923
REFERENCE	A56923
# authors	2
NOTES	
DESCRIPTION	
EDITORIAL	
REVISION	

```
#journal      Cell (1995) 81:791-800
#title       schnurri is required for Drosophila DPP signaling and encodes
```

#cross-references MUID:95292346
#accession A56923

```
##status      preliminary; nucleic acid sequence not shown; not
               compared with conceptual translation
```

```
##molecule_type mRNA
##residues 1-2529 ##label GRI
##Cross-references GB:L42311
```

#gene	FlyBase:shn
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

```
##cross-references FlyBase:FBgn0003396
KEYWORDS   DNA binding; transcription factor; zinc finger
SUMMARY    #length 2529 #molecular-weight 271637 #checksum 9689
```

Query Match	3.48;	Score 119;	DB 2;	Length 2529;
Best Local Similarity	24.68;	Pred. No. 1.40e+00;		
Matches	32;	Mismatches 55;	Indels 5;	Gaps 5;

```
Db      30 ATAAAAAAAAAAATTTVTATTPTKKR-TYRETATTATTVQRSTNKNIAAATAALMAATE 88
Qy      38 TTAATNAADLLASTTATATLVEPARITLTYYEGAPITIMONGNIIDTELLVLTLGGYKT 97
Db      89 ATASASAFATATDAT-LT-ASKAAATAATTCAGSNGSNSSSKSPSTIRDKI-GE-YP 144
Qy      98 GTTSTSVNETDAACDPMTLTFTSQDGNNHQFTTKVIKGRSDHEDISERVNGENLVGDDVV 157
Db      145 LPTVDNSHHI 154
Qy      158 LATGSODEFV 167
```

Search completed: Sat Nov 27 15:34:00 1999
Job time : 62 secs.

|||||
QY 256 NRVITDDIELLKKILATFLPDDITLKGHYDNOQNGIKRKEFLLESSPNTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKVIYDSNMHNGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 180
QY 316 VMHFSLTADRIDDDILKVIYDSNMHNGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 375
Db 181 SGTINIHDKSINLMDKNLYGTYDEEIFKASAEYKILEKMPQTTIQVDSSEKKIYSIKDFL 240
QY 376 SGTINIHDKSINLMDKNLYGTYDEEIFKASAEYKILEKMPQTTIQVDSSEKKIYSIKDFL 435
Db 241 GSEKRRGALGNLKNSSYNNKDNNELSHFATTCSDKSRPLNDIVSQKTTQSLDITSRPN 300
QY 436 GSEKRRGALGNLKNSSYNNKDNNELSHFATTCSDKSRPLNDIVSQKTTQSLDITSRPN 495
Db 301 AIEALNRFIOKIDSVWMRLDDTSGK 326
QY 496 AIEALNRFIOKIDSVWMRLDDTSGK 521

RESULT 2
ID LCRV_YERPS STANDARD: PRT: 326 AA.

AC P23994;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN DE V).
GN LCRV.
OS YERSINIA PSEUDOTUBERCULOSIS.
OC PLASMID PIB1.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-YPIII;
RX MEDLINE: 91154114.
RA BERGMAN T., HAKANSSON S., FORSBERG A., NORLANDER L., MACCELLARO A.,
BAECKMAN A., BOELIN I., WOLF-WATZ H.;
RT "Analysis of the v antigen lcrGVH-yopBD operon of Yersinia
RT pseudotuberculosis: evidence for a regulatory role of lcrH and
RT lcrV";
RT J. BACTERIOL. 173:1607-1616(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA ROGGENKAMP A., HESEEMANN J.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: INVOLVED IN CA(2+) REGULATION OF YOP EXPRESSION, WHICH
CC INCLUDES THE EXPORT PROCESS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M57893; G155458; -;
DR EMBL: X96802; E239989; -;
DR PIR: B37314; B37314.
KW PLASMID: ANTIGEN: VIRULENCE.
SQ SEQUENCE 326 AA; 37336 MW; 7D84E243 CRC32;

Query Match 61.8%; Score 2192; DB 1; Length 326;
Best Local Similarity 96.6%; Pred.No. 0.00e+00;

Matches 315; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Db 1 MIRAEQNPQHFIEDLEKVRVEQLTGSSVLELVOLVKNKIDISIKYPRKDSVEYFA 60
QY 196 MIRAEQNPQHFIEDLEKVRVEQLTGSSVLELVOLVKNKIDISIKYPRKDSVEYFA 255

Db 61 NRVITDDIELLKKILATFLPDDILKKGHYDNOQNGIKRKEFLLESSPNTQWELRAFMA 120
QY 256 NRVITDDIELLKKILATFLPDDILKKGHYDNOQNGIKRKEFLLESSPNTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKVIYDSNMHNGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 180
QY 316 VMHFSLTADRIDDDILKVIYDSNMHNGDARSKLREELAEELAEIKIYSVIOAEINKHLSS 375
Db 181 SGTINIHDKSINLMDKNLYGTYDEEIFKASAEYKILEKMPQTTIQEGETEKKIYSIKNFL 240
QY 376 SGTINIHDKSINLMDKNLYGTYDEEIFKASAEYKILEKMPQTTIQVDSSEKKIYSIKDFL 435
Db 241 ESEKRRGALGNLKNDSYNNKDNNELSHFATTCSDKSRPLNDIVSQKTTQSLDITSRPN 300
QY 436 GSEKRRGALGNLKNSSYNNKDNNELSHFATTCSDKSRPLNDIVSQKTTQSLDITSRPN 495
Db 301 AIEALNRFIOKIDSVWMRLDDTSGK 326
QY 496 AIEALNRFIOKIDSVWMRLDDTSGK 521

RESULT 3
ID CAPL_YERPE STANDARD: PRT: 170 AA.

AC P26948;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE F1 CAPSULE ANTIGEN PRECURSOR.
GN CAPL.
OS YERSINIA PESTIS.
OC PLASMID PFRA.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91099503.
RA GALIYOV E.E., SMIRNOV O.Y., KARLISHEV A.V., VOLKOVY K.I.,
RA DENEVSTUK A.I., NAZIMOV I.V., RUBTSOV K.S., ABRAMOV V.M.,
RA DALVADYANZ S.M., ZAV'YALOV V.P.;
RT "Nucleotide sequence of the Yersinia pestis gene encoding F1 antigen
RT and the primary structure of the protein. Putative T and B cell
RT epitopes";
RT FEBS LETT. 277:230-232(1990).
RN [2]
RP FBS LEFT. 277:230-232(1990).
CC -1- SUBCELLULAR LOCATION: CAPSULE.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X61996; G48621; -;
DR PIR: S13008; S13008.
KW PLASMID: SIGNAL.
FT SIGNAL 1 21
FT CHAIN 22 170
FT DOMAIN 100 150
FT FT
FT FT
SQ SEQUENCE 170 AA; 17666 MW; 354322B5 CRC32;

Query Match 31.0%; Score 1100; DB 1; Length 170;
Best Local Similarity 100.0%; Pred.No. 5.16e-166;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MKKISSVIAIALFSTIATANADITASTATATLVEPARITLYKEGAPITIMDNGINDT 60
QY 24 MKKISSVIAIALFSTIATANADITASTATATLVEPARITLYKEGAPITIMDNGINDT 83
Db 61 ELIVGTLTIGYKGTSTSVNFTDAAGDPWYLTFTSODCNHNOFTTKVIGKDSRDFDIS 120
QY 84 ELIVGTLTIGYKGTSTSVNFTDAAGDPWYLTFTSODCNHNOFTTKVIGKDSRDFDIS 143

Db 121 PKVNGENVGDVVLATGSDFFVRSIGSGKGLAAGKYDAVTYVSNQ 170
144 PKVNGENVGDVVLATGSDFFVRSIGSGKGLAAGKYDAVTYVSNQ 193
QY

RESULT 4
ID FMS3_ECOLI STANDARD; PRT; 168 AA.
AC P15488;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CS3 FIMBRIAL SUBUNIT A PRECURSOR (CS3 PILIN).
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
RC STRAIN-PB176;
RX MEDLINE; 90158116.
RA JALAJAKUWARI M.B., THOMAS C.J., HALTER R., MANNING P.A.;
RT "Genes for biosynthesis and assembly of CS3 pilin of CFA/II
enterotoxigenic Escherichia coli: novel regulation of pilus
production by bypassing an amber codon."
RT MOL. MICROBIOL. 3:1685-1695(1989).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE; 89032631.
RA BOYLAN M., SMYTH C.J., SCOTT J.R.;
RT "Nucleotide sequence of the gene encoding the major subunit of CS3
fimbriae of enterotoxigenic Escherichia coli."
RT INFECT. IMMUN. 56:3297-3300(1988).
RN [3]
RP SEQUENCE OF 23-45.
RC STRAIN-E9034A;
RX MEDLINE; 90036735.
RA HALL R.H., MANEVAL D.R. JR., COLLINS J.H., THEIBERT J.L.,
RA LEVINE M.M.;
RT "Purification and analysis of colonization factor antigen I, coli
RT enterotoxigenic Escherichia coli."
RT J. BACTERIOL. 171:6372-6374(1989).
RL -1- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X16944; G41161;
DR EMBL; M35657; G145627;
DR PIR; A34952; A34952;
DR PIR; S07904; S07904.
KW FIMBRIA; SIGNAL.
FT SIGNAL 1 22
FT CHAIN 23 168 CS3 FIMBRIAL SUBUNIT A.
FT CONFLICT 83 83 N -> S (IN REF. 2).
SQ SEQUENCE 168 AA; 17491 MW; 4D5CD89A CRC32;

Query Match 3.6%; Score 127; DB 1; Length 168;
Best Local Similarity 24.1%; Pred. No. 3.27e-02;
Matches 41; Conservative 47; Mismatches 72; Indels 10; Gaps 10;

Db 1 MLKIKYLLIGSLTSMSSSYLAAGPILTKELANVSPALDITWAPQDNLSTNG-V 59
QY 24 MKRTISSV-IATALFGTIA-TANADLTASTATATVATVETPARITLTYKEGADITIMDNGNI 81

Db 60 -SNTLVGLTLSTSDTYSIASTNSDTSKNGT-VTFAHETNNSASFAT-ISTDNANI 116
82 DTELLVGLTLIGLTKGTGTS-TSVNFTDAAGDPMYLTFTSQDNNHQFTTKVIGKDSRDF 140
QY

Db 117 TLID-KNAGNTIV-KTINGSQLPTNLPKFTTEGNEHVSNGNRANTIT 164
141 DISPKVNGENVGDVVLATGSDFFVRSIGSGKGLAAGKYDAVTYV 189
QY

RESULT 5
ID ARPA YEAST STANDARD; PRT; 489 AA.
AC P60428;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ACTIN-LIKE PROTEIN ARPA.
GN ARPA OR ACT3 OR X1081C OR J1012.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94336725.
RA HARARYA M., KARMAN A., WINTERSBERGER U.;
RT "An essential gene of Saccharomyces cerevisiae coding for an actin-
RT related protein."
RT PROC. NATL. ACAD. SCI. U.S.A. 91:8258-8262(1994).
RN [2]
RP ERRATUM.
RA HARARYA M., KARMAN A., WINTERSBERGER U.;
RT "PROC. NATL. ACAD. SCI. U.S.A. 91:10757-10757(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE; 96093911.
RA MTOGA T., SCHAEFF-GERSTENSCHLAGER I., CHALWATZIS N., BAUR A.,
RA BOLES E., FOURNIER C., SCHMITT S., VELTEN C., WILHELM N.,
RA ZIMMERMANN F.K.;
RT "Sequence analysis of a 3.1 kb fragment from the left arm of
RT Saccharomyces cerevisiae chromosome X, including putative proteins
RT with leucine zippers, a fungal Zn(II)-2-Cys6 binuclear cluster domain
RT and a putative alpha 2-SCB-alpha 2 binding site."
RL YEAST 11:681-689(1995).
RL -1- SIMILARITY: TO ACTIN. STRONG, TO OTHER ACTIN-LIKE PROTEINS
OF THE ARPA SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X75317; G436808;
DR EMBL; X83502; G929875;
DR EMBL; Z49356; G1008244;
DR PIR; S47608; S47608.
DR SGD; L00000027; ARPA.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
DR PFAM; PF00022; actin; 2.
DR HSSP; P02568; IATN.
KW STRUCTURAL PROTEIN; CYTOSKELETON.
SQ SEQUENCE 489 AA; 54831 MW; C41435B1 CRC32;

Query Match 3.5%; Score 123; DB 1; Length 489;
Best Local Similarity 25.8%; Pred. No. 9.34e-02;
Matches 16; Conservative 21; Mismatches 24; Indels 1; Gaps 1;

Db 22 GSTTITIGYSGSDPSPILPSYGVKTYADEGNKRISEOSIGIPRDYELKPIIE-NGLV 80
QY 93 GGYKTGTSTSVNFTDAAGDPMYLTFTSQDNNHQFTTKVIGKDSRDFDISPKVNGENV 152

Dh 81 ID 82
Oy 153 GD 154

RESULT 6
ID MEDB.GIALA STANDARD: PRT: 857 AA.
AC Q08014;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE MEDIAN BODY PROTEIN.
OS GIARDIA LAMBDA (GIARDIA INTESTINALIS).
OC EUGAROTIA; DIPLOMONADIDA; HEXAMITIDAE; GIARDIINAE; GIARDIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PORTLAND-1;
RX MEDLINE: 93287123.
RA MARSHALL J., HOLBERTON D.V.;
RT "Sequence and structure of a new coiled coil protein from a
microtubule bundle in Giardia."
RL J. MOL. BIOL. 231:521-530(1993).
CC -1- FUNCTION: MAY HAVE A ROLE IN IMMOBILIZING THE MICROTUBULES
BETWEEN CELL DIVISIONS.
CC -1- SUBCELLULAR LOCATION: MEDIAN BODY.
CC -1- DOMAIN: SHOWS AN ALPHA-HELICAL COILED COIL STRUCTURE (30
REPEATING HEPTADS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X64517; G312671; -
DR HSSP: P02847; 1A11.
KW CYTOSKELETON; MICROTUBULES; COILED COIL.
FT DOMAIN 120 796 COILED COIL (POTENTIAL).
SQ SEQUENCE 857 AA; 100583 MW; 5E2BBA84 CRC32;

Query Match 3.5%; Score 124; DB 1; Length 857;
Best Local Similarity 20.1%; Pred. No. 7, 20e-02;
Matches 43; Conservative 71; Mismatches 82; Indels 18; Gaps 17;

Dh 582 ERVVERIOREKNNLKKORITELQOQRTATVRETESALREKANELDGVNRQAREH 641
Oy 205 QHFIID-LEKRAVEQDTGHSVLEELVOLVAKNDIDISIKTDPKRDSEVFA-NRYITD- 261
Dh 642 EINMLDRALAESDKLRQDNRYVAMEL-TELREKVOLEKQYERARVEMLELRKAMD 700
Oy 262 DIELLK-KILAF-LPEPTILKGHYDNOGSIKRV-K-ELLESSPNTQ-WELAA-FWA 315
Dh 701 VDTLVEKORLEMLAEIKIKYNNATDQADKARLOEOLKESDKLIEFEMTMD-NRRL 759
Oy 316 VMHFSLTADRIDDDI--LKVIYDSNMHHGDARSKLREIELAELTALKIYSVIAQEIKNHL 373
Dh 760 KLOVA-ELDIKTAN-MEK-LVEEYKKLEDOUKAT 790
Oy 374 SSSGTINIHDKSINLMDKNLVG-ITD-EETFNAS 405

RESULT 7
ID FAST.YARLI STANDARD: PRT: 2076 AA.
AC P34229;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE FATTY ACID SYNTHASE, SUBUNIT BETA (EC 2.3.1.86) [CONTAINS: 3-
DEHYDROXYPALMITOYL-[ACYL-CARRIER-PROTEIN] DEHYDRATASE (EC 4.2.1.61);
DE ENOYL-[ACYL-CARRIER-PROTEIN] REDUCTASE (NADH) (EC 1.3.1.9); [ACYL-
CARRIER-PROTEIN] ACETYLTRANSFERASE (EC 2.3.1.38); [ACYL-CARRIER-

DE PROTEIN] MALONYLTRANSFERASE (EC 2.3.1.39); S-ACYL FATTY ACID SYNTHASE
DE THIOESTERASE (EC 3.1.2.14)].
GN FAS1.
OS YARROWIA LIPOLYTICA (CANDIDA LIPOLYTICA).
OC EUGAROTIA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC DIPODASCACEAE; YARROWIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CX161-1B;
RX MEDLINE: 91238709.
RA KOETIG H., ROTTNER G., BECK K.-F., SCHWEIZER M., SCHWEIZER E.;
RT "The pentafunctional FAS1 genes of Saccharomyces cerevisiae and
Yarrowia lipolytica are co-linear and considerably longer than
previously estimated."
RL MOL. GEN. GENET. 226:310-314(1991).
CC -1- FUNCTION: FATTY ACID SYNTHASE CATALYZES THE FORMATION OF
LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
CC THE BETA SUBUNIT CONTAINS DOMAINS FOR: [ACYL-CARRIER PROTEIN]
ACETYLTRANSFERASE AND MALONYLTRANSFERASE, S-ACYL FATTY ACID
SYNTHASE THIOESTERASE, ENOYL-[ACYL-CARRIER PROTEIN] REDUCTASE, AND
3-HYDROXYPALMITOYL-[ACYL-CARRIER PROTEIN] DEHYDRATASE.
CC -1- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH =
LONG-CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2N NADP(+) =
CC -1- CATALYTIC ACTIVITY: ACETYL-COA + N MALONYL-COA + 2N NADPH =
LONG-CHAIN FATTY ACID + (N+1) COA + N CO(2) + 2N NADP(+) = COA
+ ACETYL-[ACYL-CARRIER PROTEIN].
CC -1- CATALYTIC ACTIVITY: ACETYL-COA + [ACYL-CARRIER PROTEIN] = COA
+ MALONYL-[ACYL-CARRIER PROTEIN].
CC -1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYPALMITOYL-[ACYL-CARRIER PROTEIN]
= 2-HEXADECENOYL-[ACYL-CARRIER PROTEIN] + H(2)O.
CC -1- CATALYTIC ACTIVITY: ACYL-[ACYL-CARRIER PROTEIN] + NAD(+) = 2,3-
DEHYDROACYL-[ACYL-CARRIER PROTEIN] + NADH.
CC -1- CATALYTIC ACTIVITY: OLEOYL-[ACYL-CARRIER PROTEIN] + H(2)O =
ACYL-CARRIER PROTEIN + OLEATE.
CC -1- SUBUNIT: FATTY ACID SYNTHASE ARE [ALPHA(6)BETA(6)] HEXAMERS OF
TWO MULTIFUNCTIONAL SUBUNITS (ALPHA & BETA).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59690; G397855; -
DR PIR: S15989; S15999.
KW FFAA; PF00698; Acyl_transf: 1.
FT FATTY ACID BIOSYNTHESIS; MULTIFUNCTIONAL ENZYME; OXIDOREDUCTASE;
KW TRANSFERASE; HYDROLASE; LYASE; NAD; NADP.
FT DOMAIN 1 470 ACETYL TRANSFERASE.
FT DOMAIN 482 869 ENOYL REDUCTASE.
FT DOMAIN 1156 1647 DEHYDRATASE.
FT DOMAIN 1648 1869 MALONYL/PALMITOYL TRANSFERASE.
FT ACYL SITE 276 276 ACETYL TRANSFERASE (BY SIMILARITY).
FT ACYL SITE 1832 1832 MALONYLTRANSFERASE (BY SIMILARITY).
SQ SEQUENCE 2076 AA; 230223 MW; E10A5234 CRC32;

Query Match 3.5%; Score 123; DB 1; Length 2076;
Best Local Similarity 31.3%; Pred. No. 9, 34e-02;
Matches 25; Conservative 22; Mismatches 28; Indels 5; Gaps 5;

Dh 1268 TDRNRIKEFWKLMFEGQDSK-FEEDDIT-EEIIGDVTISGRAIDFVAVGNKGFAF 1325
Oy 119 TSQGNHNFQFTTKV-IGKDSRDFDISPKVNGENVGDVVLATSGQDFVATSGKGGKL 177
Dh 1326 V-GRSTAGTV-FAPMDFAI 1343
Oy 178 AAGKTYDAVVTYVTSNQEFMI 197

RESULT 8


```
RT developmental control genes."
RL CYTOGENET. CELL GENET. 66:267-271(1994).
CC -1- FUNCTION: MAY PLAY A ROLE IN DETERMINING OR MAINTAINING THE
CC IDENTITIES OF A SMALL SUBSET OF VISUAL SYSTEM NEURONS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: BRAIN. PERIPHERAL SENSORY NERVOUS SYSTEM AND
CC RETINA. IN THE ADULT NERVOUS SYSTEM BRN-3.2 PREDOMINATES IN THE
CC OPTICAL, INTERMEDIATE, AND DEEP GRAY AREAS OF THE SUPERIOR
CC COLLICULUS, THE DORSAL COLUMN OF THE MESENCEPHALIC AND PONTINE
CC CENTRAL GRAY, AND THE LATERAL INTEREDUNCULAR NUCLEUS.
CC -1- DEVELOPMENTAL STAGE: IN THE CNS IT IS SELECTIVELY EXPRESSED IN
CC POSTMITOTIC, TERMINALLY DIFFERENTIATED NEURONS.
CC -1- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS. BELONGS
CC TO CLASS-4 POU.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S68377; G545069; -
DR EMBL; S69351; G546434; -
DR MGI; MGI:102524; POU4F2.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PFM; PFM0046; homeobox; 1.
DR PFM; PFM0157; pou; 1.
DR HSSP; P10037; 1A07.
KW NUCLEAR PROTEIN; DNA-BINDING; HOMEBOX.
FT DOMAIN 1 4
FT DOMAIN 43 51 POLY-MET.
FT DOMAIN 54 67 POLY-GLY.
FT DOMAIN 69 77 POLY-SER.
FT DOMAIN 76 85 POLY-GLY.
FT DOMAIN 112 121 POLY-GLY.
FT DOMAIN 128 131 POU-IV BOX.
FT DOMAIN 157 160 POLY-HIS.
FT DOMAIN 173 184 POLY-SER.
FT DOMAIN 255 329 POLY-HIS.
FT DOMAIN 347 406 POU DOMAIN.
FT DNA_BIND 90 98 HOMEBOX.
FT CONFLICT 404 405 RRACLPTRP -> MCAFLQLQ (IN REF. 2).
FT CONFLICT 411 AA; 43172 MW; 99973657 CRC32;
SQ SEQUENCE 411 AA; 43172 MW; 99973657 CRC32;

Query Match 3.4%; Score 121; DB 1; Length 411;
Best Local Similarity 28.6%; Pred. No. 1,56e-01;
Matches 24; Conservative 22; Mismatches 35; Indels 3; Gaps 3;

DB 174 HHHHHHHHHHPPHALLGELLEHLSPGLALGAMAGPGSTVSTPAHAPHNATNPMHQA 233
OY 3 HHHHHHHHHHSSGHLDDDDKHKRISSVLAIALFGTATATANAADLASTATATLVEBA 61
DB 234 ALSMANAGLP-SHMGCSMDVAD 256
OY 62 RTLTLYKEGAPLTIMD-NGNIDTE 84

RESULT 11
AC BR3B_HUMAN STANDARD: PRT; 410 AA.
AC Q12837; Q13883; Q14987;
DT 01-NOV-1987 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE BRAIN-SPECIFIC HOMEBOX/POU DOMAIN PROTEIN 3B (BRN-3B).
GN POU4F2 OR BRN3B.
OS HOMIO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-RETINA;
RX MEDLINE; 9400832.
RA XING M., ZHOU L.-J., PENG Y., EDDY R.L., SHOWS T.B., NATHANS J.;
RT "Brn-3b: a pou domain gene expressed in a subset of retinal ganglion
RT cells.";
RL NEURON 11:689-701(1993).
RN [2]
RP SEQUENCE OF 146-410 FROM N.A.
RC TISSUE-TESTIS;
RX MEDLINE; 93324388.
RA RING C.J.A., LATCHMAN D.S.;
RT "The human Brn-3b POU transcription factor shows only limited
RT homology to the Brn-3a/RDC-1 factor outside the conserved POU
RT domain.";
RL NUCLEIC ACIDS RES. 21:2946-2946(1993).
RN [3]
RP SEQUENCE OF 275-392 FROM N.A.
RX MEDLINE; 94052142.
RA BHARGAVA A.K., LI Z., WEISSMAN S.M.;
RT "Differential expression of four members of the POU family of
RT proteins in activated and phorbol 12-myristate 13-acetate-treated
RT Jurkat T cells.";
RL PROC. NATL. ACAD. SCI. U.S.A. 90:10260-10264(1993).
CC -1- FUNCTION: MAY PLAY A ROLE IN DETERMINING OR MAINTAINING THE
CC IDENTITIES OF A SMALL SUBSET OF VISUAL SYSTEM NEURONS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: BRAIN. SEEMS TO BE SPECIFIC TO THE RETINA.
CC -1- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS. BELONGS
CC TO CLASS-4 POU.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U06233; G458391; -
DR EMBL; X71488; G312469; -
DR EMBL; L20434; G508494; -
DR MIM; 113725; -
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PFM; PFM0046; homeobox; 1.
DR PFM; PFM0157; pou; 1.
DR HSSP; P10037; 1A07.
KW NUCLEAR PROTEIN; DNA-BINDING; HOMEBOX.
FT DOMAIN 1 4
FT DOMAIN 43 51 POLY-MET.
FT DOMAIN 54 69 POLY-SER.
FT DOMAIN 71 80 POLY-GLY.
FT DOMAIN 81 84 POLY-SER.
FT DOMAIN 111 120 POLY-GLY.
FT DOMAIN 127 130 POU-IV BOX.
FT DOMAIN 156 159 POLY-HIS.
FT DOMAIN 172 183 POLY-SER.
FT DOMAIN 183 328 POLY-HIS.
FT DOMAIN 346 405 POU DOMAIN.
FT DNA_BIND 166 166 HOMEBOX.
FT CONFLICT 183 183 C -> S (IN REF. 2).
FT CONFLICT 183 183 MISSING (IN REF. 2).
SQ SEQUENCE 410 AA; 43159 MW; 48FFA52D CRC32;

Query Match 3.3%; Score 118; DB 1; Length 410;
Best Local Similarity 27.4%; Pred. No. 3,36e-01;
Matches 23; Conservative 23; Mismatches 35; Indels 3; Gaps 3;

DB 173 HHHHHHHHHHPPHALLGELLEHLSPGLALGAMAGPGAVVSTPAHAPHNATNPMHQA 232
OY 3 HHHHHHHHHHSSGHLDDDDKHKRISSVLAIALFGTATATANAADLASTATATLVEBA 61
```

```

QY 3 HHHHHHHSSGHIDDDKHKMKSSVIAI-ALFGTATANADLASTATATLVEPA 61
Db 233 ALSMAHAGLP-SHMGCSVDVAD 255
OY 62 RITLYKCAPITIMD-NGNIDTE 84

RESULT 12
ID BR11.BRARE STANDARD: PRT: 406 AA.
AC 090436; 090433;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BRAIN-SPECIFIC HOMEOBOX/POU DOMAIN PROTEIN 1.1 (BRN-1.1) (FRAGMENT).
GN BRN-1.1.
OS BRACHYDANTO RERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EUBAROTIA METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; CYPRINIFORMES; CYPRINOIDEA;
OC CYPRINIDAE; RASBORINAE; DANIO.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 96193692.
RA SAMPATH K., STUART G.W.;
RT "Developmental expression of class III and IV POU domain genes in the
RT zebrafish."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 219:565-571(1996).
CC -1- FUNCTION: MAY PLAY A ROLE IN SPECIFYING TERMINALLY DIFFERENTIATED
CC NEURONAL PHENOTYPES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: NERVOUS SYSTEM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN DEVELOPMENT AND IN THE
CC ADULT BRAIN.
CC -1- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS. BELONGS
CC TO CLASS-3 POU.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U43898; G1399384; -.
DR EMBL: U43656; G1323758; -.
DR PROSITE: PS00027; HOMEOBOX_1; 1.
DR PROSITE: PS00071; HOMEOBOX_2; 1.
DR PROSITE: PS00035; POU_1; 1.
DR PROSITE: PS00465; POU_2; 1.
DR PFAM: PF00046; homeobox; 1.
DR PFAM: PF00157; pou; 1.
DR HSP: P14859; POU.
DR KW HOMEOBOX; DNA-BINDING; TRANSCRIPTION REGULATION; NUCLEAR PROTEIN.
FT NON_TER 1 1
FT DOMAIN 228 299 POU DOMAIN.
FT DNA_BIND 317 376 HOMEOBOX.
FT DOMAIN 81 92 POLY-ALA.
FT DOMAIN 203 213 POLY-HIS.
SO SEQUENCE 406 AA; 43677 MW; DB96889B CRC32;

Query Match 3.2%; Score 114; DB 1; Length 406;
Best Local Similarity 56.5%; Pred. No. 9, 11e+01;
Matches 13; Conservative 4; Mismatches 5; Indels 1; Gaps 1;
Db 201 MDNHHHHNHQHNAG-VNSHDSH 222
OY 1 MGNHHHHHHSSGHIDDDDKH 23

RESULT 13
ID SPOT_MYCPN STANDARD: PRT: 733 AA.
AC P75386;
DT 01-NOV-1997 (REL. 35, CREATED)

```

```

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROBABLE GUANOSINE-3',5'-BIS(DIPHOSPHATE) 3'-PYROPHOSPHOHYDROLASE
DE (EC 3.1.7.2) ((PPGPP)ASE) (PENTA-PHOSPHATE GUANOSINE-3'-
DE PYROPHOSPHOHYDROLASE).
GN SPOT.
OS MYCOPLASMA PNEUMONIAE.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 29342 / M129;
RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
RA HERRMANN R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
CC -1- FUNCTION: IN EUBACTERIA PPGPP (GUANOSINE 3'-DIPHOSPHATE 5'-
CC DIPHOSPHATE) IS A MEDIATOR OF THE STRINGENT RESPONSE THAT
CC COORDINATES A VARIETY OF CELLULAR ACTIVITIES IN RESPONSE TO
CC CHANGES IN NUTRITIONAL ABUNDANCE. THIS ENZYME CATALYSES THE
CC DEGRADATION OF PPGPP INTO GDP. IT MAY ALSO BE CAPABLE OF
CC CATALYZING THE SYNTHESIS PPGPP (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: GUANOSINE-3',5'-BIS(DIPHOSPHATE) + H(2)O =
CC GUANOSINE-5'-DIPHOSPHATE + PYROPHOSPHATE.
CC -1- COFACTOR: MANGANESE (BY SIMILARITY).
CC -1- SIMILARITY: HIGH TO RELA PROTEIN WHICH CATALYSES PPGPP FROM
CC ATP AND GTP.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000043; G1674128; -.
DR KW HYDROLASE; MANGANESE.
SO SEQUENCE 733 AA; 86403 MW; 2654A136 CRC32;

Query Match 3.2%; Score 113; DB 1; Length 733;
Best Local Similarity 27.0%; Pred. No. 1, 1e+00;
Matches 33; Conservative 34; Mismatches 45; Indels 10; Gaps 9;
Db 387 SRDEFKRIFGTDIKVNVSDNEPNIYKLVNESKLLDIAYELYPKOF-NKLEKIKLD 445
OY 162 SQDFVRSIGSGKGLA-AGKY-IDAVT-VTV-SNOERMIAEQNHFIEDLEKYRVE 217
Db 446 GVEVMSFDVTANEMVIEFCGKTNKRLRWLRMNNHVERERY-KDNLKRAVYSE 504
OY 218 -QLTGHSVLEBLV-OLVYDKNKIDISIKYDPKRDSEVFANRYITDIELKRLAYF- 273
Db 505 LP 506
OY 274 LP 275

RESULT 14
ID NIPL YEAST STANDARD: PRT: 812 AA.
AC P32497;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE NUCLEAR TRANSPORT PROTEIN NIPL.
GN NIPL OR YMR309C OR YMR924.01C OR YMR952.11C.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUBAROTIA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC SACCHAROMYCETACEAE; SACCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93066237.

```

RA GU 2., MOERSCHHELL R.P., SHERMAN F., GOLDFARB D.S.;
RT "NIP1, a gene required for nuclear transport in yeast."
RL PROC. NATL. ACAD. SCI. U.S.A. 89:10355-10359(1992).
RN [2]
RP SEQUENCE OF 1-602 FROM N.A.
RC STRAIN-S288C / AB972.
RA CHUCHER C.M., LOUIS E.J., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 571-812 FROM N.A.
RC STRAIN-S288C / AB972.
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC - FUNCTION: INVOLVED IN TRANSPORT OF PROTEINS TO THE NUCLEUS;
CC PROBABLY VIA ITS SERINE-RICH ACIDIC N-TERMINAL.
CC - SUBUNIT: MONOMER.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC; MAINLY.
CC - SIMILARITY: TO S.POMBE SPAC1E11.01C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L02899; -: NOT_ANNOTATED_CDS.
DR EMBL: 254141; G984682; -
DR EMBL: 249212; G798951; -
DR PIR: A46417; A46417.
DR SGD: L0001252; NIP1.
KW TRANSPORT; PROTEIN TRANSPORT.
FT DOMAIN 15 147 ASP/GLU-RICH (ACIDIC).
FT 15 36 SER-RICH
FT CONFLICT 111 111 V -> D (IN REF. 1).
FT CONFLICT 583 583 Q -> H (IN REF. 1).
FT CONFLICT 641 641 K -> N (IN REF. 1).
FT CONFLICT 643 643 R -> N (IN REF. 1).
SQ SEQUENCE 812 AA; 93204 MW; 47993DCE CRC32;

Query Match 3.2%; Score 115; DB 1; Length 812;
Best Local Similarity 29.2%; Pred. No. 7.11e-01;
Matches 31; Conservative 28; Mismatches 39; Indels 8; Gaps 8;

DB 700 WALLPMETVLSLT-ERVQVESLKTFFSFRFYSFSVAKL-AELFDL-PENKVEVL 756
QY 308 WEIRAMAMVHSLTADRIDDLKIVDSMNH-HGD-ARSKLRELAELAEELKIVSYI 365
DB 757 QSVIA-LEIIPAKLN-DEKTIPEVKEGDEITKLEAMVNLNREYKI 800
QY 366 QAEINKHLSSSGTINHDKSINLMCK-NLYGYTDEEIFKASAEYKI 410

RESULT 15
ID HA34_BRELIC STANDARD; PRI; 173 AA.
AC Q99074;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE HAM34 PROTEIN.
GN HAM34.
OS BREMIA LACTUCAE (LETTUCE DOWNY MILDEN).
OC EUKARYOTA; STRAMENOPHYTES; COMYCETES; PERONOSPORALES; PERONOSPORALES;
OC BREMIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REGEL; TISSUE=SPORE;
RX MEDLINE: 92033077.
RA "DELSON H.S., MICHELMORE R.W.;
RT "Highly abundant and stage-specific mRNAs in the obligate pathogen
RT Bremlia lactucaae."
RL MOL. PLANT MICROBE INTERACT. 3:225-232(1990).

CC - FUNCTION: COULD BE A STRUCTURAL PROTEIN REQUIRED FOR THE
CC INFECTION PROCESS OF B. LACTUCAE.
CC - TISSUE SPECIFICITY: GERMINATING SPORES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16984; G2488; -
KW SPOULATION; STRUCTURAL PROTEIN.
SQ SEQUENCE 173 AA; 16478 MW; 43864282 CRC32;

Query Match 3.1%; Score 109; DB 1; Length 173;
Best Local Similarity 24.3%; Pred. No. 3.06e-00;
Matches 28; Conservative 28; Mismatches 58; Indels 1; Gaps 1;

DB 1 MKFSQILVLAIAVAIAISADAPAPATPPDPAATATPAATTTTTTTTTPADAGTASTEQ 60
QY 25 KRISVIAIALFGTITATNADLTASTATATVAPATITLTKGCAITIMDNQIDTE 84
DB 61 TTAGEPAAGATNGTTTTPPADGTQTATAPLDATATESSASGEMTPV-GTDTSD 114
QY 85 LLVGTLLTGGYKTTGTTSTSVNETDAAGDPWLTFTSDGNNHQTFTKVIKDSRD 139

Search completed: Sat Nov 27 15:35:00 1999
Job time : 43 secs.


```
QY 436 GSEKRTGALGNLKNYSYKNNELSHFATTCDSKSRPLNDLVSOXTTQLSDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLDDTSCK 326
QY 496 AIEALNRFIOKYDSVMORLDDTSCK 521

RESULT 2
ID 087495 PRELIMINARY: PRT: 324 AA.
AC 087495;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
LN 1.
LCRV.
GN ICRV.
OS YERSINIA ENTEROCOLITICA.
OG PLASMID PYV.
OC BACTERIA; PROTOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-W22703;
RA TRIARTE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
RT "Complete sequence of the yersinia enterocolitica pyv plasmid.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF080155; G360358; -.
KW PLASMID.
SQ SEQUENCE 324 AA; 37296 MW; 7C9A652F CRC32;

Query Match 60.9%; Score 2159; DB 2; Length 324;
Best Local Similarity 95.4%; Pred. No. 0.00e+00;
Matches 308; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Db 1 MIRAYEQNPQHFIEDLEKAVVQDLTGHGSSVLELVQVYKDKIDISIKYDPKDSSEVFA 60
QY 196 MIRAYEQNPQHFIEDLEKAVVQDLTGHGSSVLELVQVYKDKIDISIKYDPKDSSEVFA 255
Db 61 DRYITDDIELKKILAYFLPEDAILKGGHYDNLQNGIKRKYEFLESSPTQWELRAFMA 120
QY 256 NRYITDDIELKKILAYFLPEDITLKGHYDNLQNGIKRKYEFLESSPTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTYSVQAEINKHLSN 180
QY 316 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTYSVQAEINKHLS 375
Db 181 SDTINIHDKSINIMDKNLGYTDEEIFKASAEYKILKMPQTTIKEGETEKKIYSINFL 240
QY 376 SGTINIHDKSINIMDKNLGYTDEEIFKASAEYKILKMPQTTIQVGSSEKKIYSIKDFL 435
Db 241 ESENKRTGALGNLKNYSYKNNELSHFATTCDSKSRPLNDLVSOXTTQLSDITSRFS 300
QY 436 GSEKRTGALGNLKNYSYKNNELSHFATTCDSKSRPLNDLVSOXTTQLSDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLDDT 323
QY 496 AIEALNRFIOKYDSVMORLDDT 518

RESULT 3
ID 056880 PRELIMINARY: PRT: 325 AA.
AC 056880;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
LN 1.
LCRV.
GN ICRV.
OS YERSINIA ENTEROCOLITICA.
OG BACTERIA; PROTOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-8081;
RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
RX MEDLINE: 97162308.
SQ SEQUENCE 324 AA; 37162 MW; 56842D0A CRC32;
```

```
RT "Passive immunity to infection with Yersinia spp. mediated by
anti-recombinant V antigen is dependent on polymorphism of V
antigen.";
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96798; E239984; -.
FT NON TER 325
SQ SEQUENCE 325 AA; 37260 MW; D6E88446 CRC32;

Query Match 60.1%; Score 2132; DB 2; Length 325;
Best Local Similarity 94.4%; Pred. No. 0.00e+00;
Matches 305; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

Db 1 MIRAYEQNPQHFIEDLEKAVVQDLTGHGSSVLELVQVYKDKIDISIKYDPKDSSEVFA 60
QY 196 MIRAYEQNPQHFIEDLEKAVVQDLTGHGSSVLELVQVYKDKIDISIKYDPKDSSEVFA 255
Db 61 DRYITDDIELKKILAYFLPEDAILKGGHYDNLQNGIKRKYEFLESSPTQWELRAFMA 120
QY 256 NRYITDDIELKKILAYFLPEDITLKGHYDNLQNGIKRKYEFLESSPTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTYSVQAEINKHLSN 180
QY 316 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTYSVQAEINKHLS 375
Db 181 SGTINIHDKSINIMDKNLGYTDEEIFKASAEYKILKMSQTTIKEGETEKKIYSIKNFL 240
QY 376 SGTINIHDKSINIMDKNLGYTDEEIFKASAEYKILKMPQTTIQVGSSEKKIYSIKDFL 435
Db 241 ESENKRTGALGNLKNYSYKNNELSHFATTCDSKSRPLNDLVSOXTTQLSDITSRFS 300
QY 436 GSEKRTGALGNLKNYSYKNNELSHFATTCDSKSRPLNDLVSOXTTQLSDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLDDT 323
QY 496 AIEALNRFIOKYDSVMORLDDT 518

RESULT 4
ID 056882 PRELIMINARY: PRT: 324 AA.
AC 056882;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
LN 1.
LCRV.
GN ICRV.
OS YERSINIA ENTEROCOLITICA.
OG BACTERIA; PROTOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y-96-P;
RX MEDLINE: 97162308.
RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HEESEMANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by
anti-recombinant V antigen is dependent on polymorphism of V
antigen.";
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96801; E239988; -.
SQ SEQUENCE 324 AA; 37162 MW; 56842D0A CRC32;

Query Match 60.0%; Score 2128; DB 2; Length 324;
Best Local Similarity 93.8%; Pred. No. 0.00e+00;
Matches 303; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

Db 1 MIRAYEQNPQHFIEDLEKAVVQDLTGHGSSVLELVQVYKDKIDISIKYDPKDSSEVFA 60
QY 196 MIRAYEQNPQHFIEDLEKAVVQDLTGHGSSVLELVQVYKDKIDISIKYDPKDSSEVFA 255
Db 61 DRYITDDIELKKILAYFLPEDAILKGGHYDNLQNGIKRKYEFLESSPTQWELRAFMA 120
QY 256 NRYITDDIELKKILAYFLPEDITLKGHYDNLQNGIKRKYEFLESSPTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTYSVQAEINKHLSN 180
```



```

OY 316 VMHSLTADRIDDDILKVIYDSNMHNGDARSKLRELAELTAEIKIYVIOAEINKHLS 375
DB 181 SDTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIKEGETEKIYSIKNFL 240
OY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSEKKIYSIKNFL 435
DB 241 ESENKRTGALGNLKDYSYNKDNNELSHLGTTCSDKSRPLNDVLSQKTTQSLDITSRFS 300
OY 436 GSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVLSQKTTQSLDITSRFS 495
DB 301 AIEALNRFIOKYDSVMORLLDDT 323
OY 496 AIEALNRFIOKYDSVMORLLDDT 518

RESULT 5
ID 056879 PRELIMINARY: PRT: 324 AA.
AC 056879;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE V ANTIGEN
OS YERSINIA ENTEROCOLITICA.
OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97162308.
RA "ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESSEMANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by
RT anti-recombinant V antigen is dependent on polymorphism of V
RT antigen."
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96800; E239987;
SQ SEQUENCE 324 AA: 37195 MW: EICB300F CRC32;

Query Match 59.9%; Score 2125; DB 2; Length 324;
Best Local Similarity 93.8%; Pred. No. 0.00e+00;
Matches 303; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

DB 1 MIRAEQNPQHFIDLEKVRVEQLTGHSVLEELVQLVODKKIDISIKYDPKDSVFA 60
OY 196 MIRAEQNPQHFIDLEKVRVEQLTGHSVLEELVQLVODKKIDISIKYDPKDSVFA 255
DB 61 DRVITDDIELLKKILAYFLPEDAILKGHYDQNGIKRVEKEFLESSPNTQWELRAFMA 120
OY 256 NRVITDDIELLKKILAYFLPEDAILKGHYDQNGIKRVEKEFLESSPNTQWELRAFMA 315
DB 121 VMHSLTADRIDDDILKVIYDSNMHNGDARSKLRELAELTAEIKIYVIOAEINKHLSN 180
OY 316 VMHSLTADRIDDDILKVIYDSNMHNGDARSKLRELAELTAEIKIYVIOAEINKHLS 375
DB 181 SDTINIHDKSINLMDKNLYGTNEEIFKASAEYKILEKMSQTTIKEGETEKIYSIKNFL 240
OY 376 SGTINIHDKSINLMDKNLYGTNEEIFKASAEYKILEKMPOTTIOVDSEKKIYSIKNFL 435
DB 241 ESENKRTGALGNLKDYSYNKDNNELSHLGTTCSDKSRPLNDVLSQKTTQSLDITSRFS 300
OY 436 GSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVLSQKTTQSLDITSRFS 495
DB 301 AIEALNRFIOKYDSVMORLLDDT 323
OY 496 AIEALNRFIOKYDSVMORLLDDT 518

RESULT 6
ID 056846 PRELIMINARY: PRT: 324 AA.
AC 056846;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE V ANTIGEN.
OS YERSINIA ENTEROCOLITICA.

```

```

OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y-108-P;
RX MEDLINE: 97162308.
RA "ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESSEMANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by
RT anti-recombinant V antigen is dependent on polymorphism of V
RT antigen."
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96796; E239985;
SQ SEQUENCE 324 AA: 37194 MW: 5BD357FF CRC32;

Query Match 59.9%; Score 2124; DB 2; Length 324;
Best Local Similarity 94.1%; Pred. No. 0.00e+00;
Matches 304; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

DB 1 MIRAEQNPQHFIDLEKVRVEQLTGHSVLEELVQLVODKKIDISIKYDPKDSVFA 60
OY 196 MIRAEQNPQHFIDLEKVRVEQLTGHSVLEELVQLVODKKIDISIKYDPKDSVFA 255
DB 61 NRVITDDIELLKKILAYFLPEDAILKGHYDQNGIKRVEKEFLESSPNTQWELRAFMA 120
OY 256 NRVITDDIELLKKILAYFLPEDAILKGHYDQNGIKRVEKEFLESSPNTQWELRAFMA 315
DB 121 VMHSLTADRIDDDILKVIYDSNMHNGDARSKLRELAELTAEIKIYVIOAEINKHLSN 180
OY 316 VMHSLTADRIDDDILKVIYDSNMHNGDARSKLRELAELTAEIKIYVIOAEINKHLS 375
DB 181 SDTINIHDKSINLMDKNLYGTNEEIFKASAEYKILEKMSQTTIKEGETEKIYSIKNFL 240
OY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSEKKIYSIKNFL 435
DB 241 ESENKRTGALGNLKDYSYNKDNNELSHLGTTCSDKSRPLNDVLSQKTTQSLDITSRFS 300
OY 436 GSENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDVLSQKTTQSLDITSRFS 495
DB 301 AIEALNRFIOKYDSVMORLLDDT 323
OY 496 AIEALNRFIOKYDSVMORLLDDT 518

RESULT 7
ID 056853 PRELIMINARY: PRT: 334 AA.
AC 056853;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE V ANTIGEN (FRAGMENT).
OS YERSINIA ENTEROCOLITICA.
OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WA-314;
RX MEDLINE: 97162308.
RA "ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESSEMANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by
RT anti-recombinant V antigen is dependent on polymorphism of V
RT antigen."
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96796; E239982;
FT NON_TER 334 334
SQ SEQUENCE 334 AA: 38241 MW: 71229360 CRC32;

Query Match 58.5%; Score 2075; DB 2; Length 334;
Best Local Similarity 93.1%; Pred. No. 0.00e+00;
Matches 309; Conservative 10; Mismatches 4; Indels 9; Gaps 3;

DB 1 MIRAEQNPQHFIDLEKVRVEQLTGHSVLEELVQLVODKKIDISIKYDPKDSVFA 60
OY 196 MIRAEQNPQHFIDLEKVRVEQLTGHSVLEELVQLVODKKIDISIKYDPKDSVFA 255

```

Db	61	ERVATIDIELLKRLIAFLPEDALILKGGHYNOQONQIRKKEFELLESSPTMOWELRAFMA	120
Qy	256	NRVITDIEILKRLIAFLPEDTILKGGHYNOQONQIRKKEFELLESSPTMOWELRAFMA	315
Db	121	VMHSLTADRIDDILKVIYDSNMHHGDANSKLEELAEELTAELKITYSVYQAEINHLSS	180
Qy	316	VMHSLTADRIDDILKVIYDSNMHHGDARSKLEELAEELTAELKITYSVYQAEINHLSS	375
Db	181	SGTINIEHRSINLMDKNLYGTDEEIFKASAEKYILKMPQTTIKDDELHEVGYIAEK	240
Qy	376	SGTINIEHRSINLMDKNLYGTDEEIFKASAEKYILKMPQTTI-----Q---V-DOSEK	426
Db	241	QIVSIKNEFLESENKRTGALGNLKOSIYINKDNNELSHFPAKSDKSRPLNDVLSQKTTOL	300
Qy	427	KIVSIKNEFLESENKRTGALGNLKOSIYINKDNNELSHFPAKSDKSRPLNDVLSQKTTOL	486
Db	301	SDITSRFNSAIEALNRFQOKDYDMORLDDT	332
Qy	487	SDITSRFNSAIEALNRFQOKDYDMORLDDT	518

RESULT	8	PRELIMINARY:	PRF:	334 AA.
ID	056896			
AC	056896:			
DT	01-NOV-1996 (TREMBLRL. 01.			
DT	01-NOV-1996 (TREMBLRL. 01.			
DT	01-NOV-1998 (TREMBLRL. 08.			
DE	V ANTIGEN (FRAGMENT).			
OS	YERSINIA ENTEROCOLITICA.			
OC	BACTERIA: PROTEOBACTERIA: GAMMA SUBDIVISION; ENTEROBACTERIACEAE;			
OC	YERSINIA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-NCTC:			
RC	MEDLINE: 97162308			
RA	ROGGENKAMP A., GELGER A.M., LEITRITZ L., KESSLER A., HESSEMANN J.			
RT	"Passive immunity to infection with <i>Yersinia</i> sp. mediated by			
RT	anti-recombinant V antigen is dependent on polymorphism of V			
RT	antigen."			
RL	INFECT. IMMUN. 65:446-451(1997).			
DR	EMBL: X60797: E239983: -.			
FT	NON_TER			
FT	334			
FT	334			
FT	334 AA:			
FT	38227 MW:			
FT	F23FDFD3 CRC32:			
FT	SEQUENCE			

Query Match	58.5%;	Score 2074;	DB 2;	Length 334;
Best Local Similarity	92.8%;	Pred. No. 0.00e+00;		
Matches	308;	Conservative	11;	Mismatches 4; Indels 9; Gaps 3

Db	1	MIRAEYNPOQHIEDLEKVRVQDLTGSSSVLEETVQLVYDKRIDISIKDPKDSVFA	60
Qy	196	MIRAEYNPOQHIEDLEKVRVQDLTGSSSVLEETVQLVYDKRIDISIKDPKDSVFA	255
Db	61	ERVITDDIELLKILAYFELPDAILKGGHYDNQNGIKRKEFLESSPTQWELRAFMA	120
Qy	256	NRVITDDIELLKILAYFELPEDITLKGGHYDNQNGIKRKEFLESSPTQWELRAFMA	315
Db	121	VMHESLADRIDDDILKAYIVDSNNHHGDARSKLEFELAEFLTAEIKIYSVTQAEINKHLSS	180
Qy	316	VMHESLADRIDDDILKAYIVDSNNHHGDARSKLEFELAEFLTAEIKIYSVTQAEINKHLSS	375
Db	181	SGTINIHHSKINIMDKNLXYGTDEEIFKASAEYIILKKMQTTIKDELHEVGYIAEAK	240
Qy	376	SGTINIHHSKINIMDKNLXYGTDEEIFKASAEYIILKKMQTTI-----V-DSKE	426
Db	241	QIVSIKNDLESENKRTGALGNLKDSYSYNNKNNNELSHFATPACSPKRLNDIVSOAKTOL	300
Qy	427	KIVSIKNDLESENKRTGALGNLKDSYSYNNKNNNELSHFATPACSPKRLNDIVSOAKTOL	486
Db	301	SDITSRENSAIEALNRFIOKDYDWORLLDDT	332
Qy	487	SDITSRENSAIEALNRFIOKDYDWORLLDDT	518

RESULT	9		
ID	030527	PRELIMINARY;	PRT; 294 AA.
AC	030527;		
DT	01-JAN-1998	(TREMBLREL. 05, CREATED)	
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)	
DT	01-NOV-1998	(TREMBLREL. 08, LAST ANNOTATION UPDATE)	
DE	PCRV.		
GN	PCRV.		
OS	PSEUDOMONAS AERUGINOSA.		
OC	BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;		
OC	PSEUDOMONAS.		
OC	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN-388;		
RC	MEDLINE; 98037517.		
RA	YAHIR T.L., MENDE-MUELLER L.M., FRIESE M.B., FRANK D.W.;		
RT	"Identification of type III secreted products of the Pseudomonas		
RT	aeruginosa exoenzyme S regulon.";		
RL	J. BACTERIOL. 179:7165-7168(1997).		
DR	EMBL: AF010149; G2459968, ...		
SO	SEQUENCE 294 AA; 32283 MW; 0FAD2082 CRC32;		

	Query Match	18.4%	Score 653;	DB 2;	length 294;
	Best Local Similarity	46.3%	Pred. No. 1,64e-79;		
	Matches	105;	Conservative	53;	Mismatches 57; Indels 12; Gaps 8
Db	74	GLEVLRYLQARPPGACQMDLREFLVSATYPSLHG-RLDEVDYIGYKVKVLTQDQGRKALL	132		
QY	292	GIRKRRFLESS--PNNWELRLARMAVHHSFTLADRIDDDILIKYIVSMNHNGDARSKLR	349		
Db	133	DELKALFAELRYEVSVIOSIQAALSAKOGIRIDAGIDVLDPITYGAVGPPRKMSPEY	192		
QY	350	EELAEFLRAELKITYSVIOAEIKHKLHSSSGTIIHMKSLINLMDKNLYGTT-DEELTKAAAEY	408		
Db	193	ALLSNL-DT-F--SG---KL-SINDPLSGSPQSGELKGSIDEXPEEKDNNPVGNFATY	244		
QY	409	KILKMPQTIQVGGSEKKIYISIDPLFGSEKRRFGALGNLKNSSYNNDNNELSHPATTC	468		
Db	245	SDSRPRLNDKYNKETTLLINDSSRYNAVENLNFQIKYDSVYLDIL	291		
QY	469	SDKSRPLNDLVQSKTQTQLDSTLSTFNSAIEKLNLFQIKYDSVYWRLL	515		

RESULT	ID	PRELIMINARY:	PR:	168 AA.
AC	016256;			
DT	01-NOV-1996 (TREMBLREL. 01, CREATED)			
DT	01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)			
DT	01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)			
DE	WT1-WILMS' TUMOR SUPPRESSOR PROTEIN.			
GN	WT1.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES.			
OC	CATARRHINI; HOMINIDAE; HOMO.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 95166649.			
RA	HAMILTON T.B., BARILDA K.C., ROMANIK P.J.;			
RT	"High affinity binding sites for the Wt1's' tumour suppressor protein			
RT	WT1."			
RL	NUCLEIC ACIDS RES. 23:277-284(1995).			
EMBL	S75264: G896247; -			
DR	PROSITE: PS00086; ZINC_FINGER_C2H2; 4.			
DR	PFAM: PF00086; Z1-C2H2; 4.			
QW	ZINC-FINGER; METAL-BINDING; DNA-BINDING; SEQUENCE 168 AA; 20165 MW; E86026CC CRC32;			

Query Match	4.7%	Score 168;	DB 4;	Length 168;
Best Local Similarity	66.7%	Pred. No. 1.19e-06;		
Matches	20;	Conservative	7;	Mismatches 1;
				Indels 2;
				Gaps 1
Db	1	MGNHHHHHHHSSGHIEG--RHNRPPGV	28	

```

OY 1 MGHNNNNNNHHSSGHITDDDDKXMKKISSV 30

RESULT 11
ID 09996 PRELIMINARY: PRT: 1620 AA.
AC 09996:
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE YOFIAO (FRAGMENT).
OS HOMO SAPIENS (HUMAN).
OC EURKRYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA WU X., GRAVES T., BRADSHAW H.;
RT "The sequence of Homo sapiens BAC clone RG293F11."
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA WATERSTON R.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA WATERSTON R.;
RL SUBMITTED (SEP-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AC000066; G3645944;
FT NON-TER 1
FT SEQUENCE 1620 AA; 186793 MW; 2A1E553C CRC32;

Query Match 3.8%; Score 133; DB 4; Length 1620;
Best Local Similarity 21.9%; Pred. No. 1.69e-02;
Matches 57; Conservative 79; Mismatches 102; Indels 22; Gaps 20;

DB 247 AQQILITHOQOOLEDDHLEDYQKK-EDFTMQISFLOEKIKYEMEDQKVENSK-EE 304
OY 191 SNOEFMIRAYE-QNPOHFIEDELEKRVVEQLTGHGSSVLEL-V-QLVKKNIDISIKYDP 247
DB 305 IOEKETIEELNTKIIIEEKKTELKDKLTITADKLGLGEOIVQKNOEIKMKLELTMS 364
OY 248 RKOSEFARVITDDIELLKILAYF--LPEDTILKGHYDQLO-N-GIKRVK-EFILES 302
DB 365 KOKEROSSEIKOLMGTYE-ELOKRNKDSOFTDVOIMEQETQKLEQLRAELDEMVG 423
OY 303 SPNT-QW-ELRAFNAVHFSL-TADRIDDLIKYIVDSMNHGDAR-SKLRELAELTA 357
DB 424 Q-QIVQMKQELIRQMAQMEKTRHG-E-MENALRSYNTIV-NED-QIKLMAVINE 478
OY 358 ELKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGTDEIFKASAEKYLE-KMPQ 416
DB 479 LNIKLODPTNSQEKLEELG 498
OY 417 TTIOVDSSEKIVSIKDFLG 436

RESULT 12
ID 014869 PRELIMINARY: PRT: 1642 AA.
AC 014869:
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE YOFIAO
OS HOMO SAPIENS (HUMAN).
OC EURKRYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA LIN J.M., WYSZYNSKI M., MADHAVAN R., SEALOCK R., KIM J.U., SHENG M.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF026245; G2623068;
FT NON-TER 1
FT SEQUENCE 1642 AA; 191023 MW; 99E0E48C CRC32;

```

```

Query Match 3.8%; Score 133; DB 4; Length 1642;
Best Local Similarity 21.9%; Pred. No. 1.69e-02;
Matches 57; Conservative 79; Mismatches 102; Indels 22; Gaps 20;

DB 278 AQQILITHOQOOLEDDHLEDYQKK-EDFTMQISFLOEKIKYEMEDQKVENSK-EE 335
OY 191 SNOEFMIRAYE-QNPOHFIEDELEKRVVEQLTGHGSSVLEL-V-QLVKKNIDISIKYDP 247
DB 336 IOEKETIEELNTKIIIEEKKTELKDKLTITADKLGLGEOIVQKNOEIKMKLELTMS 395
OY 248 RKOSEFARVITDDIELLKILAYF--LPEDTILKGHYDQLO-N-GIKRVK-EFILES 302
DB 396 KOKEROSSEIKOLMGTYE-ELOKRNKDSOFTDVOIMEQETQKLEQLRAELDEMVG 454
OY 303 SPNT-QW-ELRAFNAVHFSL-TADRIDDLIKYIVDSMNHGDAR-SKLRELAELTA 357
DB 455 Q-QIVQMKQELIRQMAQMEKTRHG-E-MENALRSYNTIV-NED-QIKLMAVINE 509
OY 358 ELKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGTDEIFKASAEKYLE-KMPQ 416
DB 510 LNIKLODPTNSQEKLEELG 529
OY 417 TTIOVDSSEKIVSIKDFLG 436

RESULT 13
ID 060335 PRELIMINARY: PRT: 882 AA.
AC 060335:
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE KIA0594 PROTEIN (FRAGMENT).
GN KIA0594
OS HOMO SAPIENS (HUMAN).
OC EURKRYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-BRAIN;
RX MEDLINE: 98290545.
RA NAGASE T., ISHIKAWA K., MIYAJIMA N., TANAKA A., KOTANI H., NOMURA N.,
RA OHARA O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA RES. 5:31-39(1998).
DR EMBL: AB011166; D1026450;
FT NON-TER 1
FT SEQUENCE 882 AA; 104906 MW; 84CC9E80 CRC32;

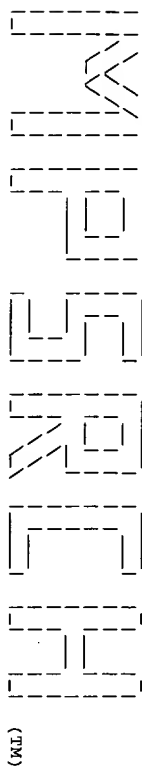
Query Match 3.6%; Score 128; DB 4; Length 882;
Best Local Similarity 25.4%; Pred. No. 6.05e-02;
Matches 45; Conservative 41; Mismatches 81; Indels 10; Gaps 10;

DB 431 EQRHLEQLKEIHRKLAVDGSLALRPTSKYLEHKN-ELROKKELLEKTKK-RQL 488
OY 343 DAKSKRELAELTAELK-TYS-VIOA-EINKHLSSSGTINIHDKSINLMDKNLYGYTDE 399
DB 489 EQ-KISSKLGSLKIMEODTCNLEEEERKASTKIKEINQAKLYVETLYLICTSLHTO 547
OY 400 ELFKASAEKILEKMPQTTIOVDSSEKIVS-INKFLSEKNRTGALGLU-KNSYSYND 457
DB 548 KYDLIQNTTVISERNKLESYMAA-SSQLRLTEQHFLELDENRORLLQCKEELMKR 603
OY 458 NNELS-HEATCSKSRPLNDLVSOKTQLOSDITSRFNSAIBALNRFIQKYSVQR 513

RESULT 14
ID 099247 PRELIMINARY: PRT: 1116 AA.
AC 099247:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

```

DT 01-NOV-1998 (TREMBLREL, 08, LAST ANNOTATION UPDATE)
DE ORF YOLO87C FROM CHROMOSOME XV.
OS SACCCHAROMYCES CREVISIAE (BAKER'S YEAST).
OC EUAROTIA, FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FY1679;
RX MEDLINE: 96021609.
RA ZUMSTEIN E., PEARSON B.M., KALOGEROPOULOS A., SCHWEITZER M.;
RT "A 29,425 kb segment on the left arm of yeast chromosome XV contains
RT more than twice as many unknown as known open reading frames.";
RL YEAST 11:975-986(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA ZUMSTEIN E., PEARSON B.M., KALOGEROPOULOS A., SCHWEITZER M.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA MIRS;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: X83121; G600471; -;
DR EMBL: Z74829; E251885; -;
SO SEQUENCE 1116 AA; 125361 MW; 099CFCF031 CRC32;



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch.p protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Nov 27 15:30:18 1999; Maspar time 20.48 Seconds

Tabular output not generated. 541.086 Million cell updates/sec

Title: >US-08-699-716A-2

Description: (1-521) from US08699716A.pep

Perfect Score: 3546

Sequence: 1 MGNHHHHHHHSSGHIDDD.....RFIOKYDSVMQRLDDTSCK 521

Scoring table: PAM 150

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-geneseq35

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics: Mean 35.735; Variance 201.794; scale 0.177

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	3294	92.9	501 19	W01044	Y. pestis FI/V antige	2.07e-236
2	3166	89.3	480 19	W01045	Y. pestis FI/V antige	1.17e-226
3	2242	63.2	329 19	W01040	Y. pestis V antigen.	2.16e-156
4	2242	63.2	329 15	R79961	Partial lcrV (V antig	6.15e-156
5	2236	63.0	329 19	W01041	Y. pestis V antigen.	1.04e-155
6	2233	63.0	329 15	R79962	Partial lcrV (V antig	3.31e-70
7	1100	31.0	170 19	W01043	Y. pestis FI antigen	3.31e-70
8	1100	31.0	170 14	R76528	Yersinia pestis cafI	3.31e-70
9	1100	31.0	170 34	M59783	Amino acid sequence o	3.31e-70
10	1100	31.0	170 34	M59782	Nucleotide sequence o	1.42e-60
11	971	27.4	149 34	M59788	Amino acid sequence o	1.42e-60
12	971	27.4	150 34	M59787	Y. pestis FI antigen.	2.38e-60
13	968	27.3	151 19	W01042	Yersinia pestis cafI	3.36e-60
14	968	27.3	151 14	R76526	Yersinia pestis cafI	3.36e-60
15	966	27.2	151 14	R76527	Amino acid sequence o	7.78e-53
16	867	24.5	192 34	M59785		

17	738	20.8	171 34	M59786	Amino acid sequence o	2.77e-43
18	190	5.4	24 23	W19779	Polyhistidine-enterok	3.17e-04
19	189	5.3	473 35	M68400	Clostridium botulinum	3.69e-04
20	172	4.9	472 35	M68398	Clostridium botulinum	4.61e-03
21	172	4.9	472 35	M68394	Clostridium botulinum	4.61e-03
22	157	4.4	451 35	M68398	Clostridium botulinum	4.11e-02
23	152	4.3	448 35	M68399	Clostridium botulinum	8.45e-02
24	154	4.3	462 35	M68397	Clostridium botulinum	6.34e-02
25	150	4.2	462 17	R95009	Type A neurotoxin C f	1.12e-01
26	150	4.2	462 35	M68380	Clostridium botulinum	1.12e-01
27	146	4.1	53 28	M36986	HEFT peptide.	1.99e-01
28	146	4.1	54 28	M36990	HEFT peptide.	1.99e-01
29	146	4.1	323 34	M71185	FP505 protein contain	1.99e-01
30	147	4.1	423 34	M68098	Chlamydomonas reinhar	1.73e-01
31	146	4.1	451 35	M68395	Clostridium botulinum	1.99e-01
32	146	4.1	452 35	M68396	Clostridium botulinum	1.99e-01
33	135	3.8	21 23	M23647	Recombinant squirrel	9.37e-01
34	135	3.8	21 23	M23650	Recombinant squirrel	9.37e-01
35	135	3.8	21 23	M23653	Recombinant squirrel	9.37e-01
36	123	3.5	168 27	M38342	E. coli colonisation	4.90e+00
37	120	3.4	290 35	M71474	Helicobacter polypept	7.36e+00
38	116	3.3	254 28	M37689	Protein sequence of c	1.28e+01
39	110	3.1	835 17	R96206	Invasin protein.	2.79e+01
40	106	3.0	174 17	R88503	Borrelia burgdorferi	4.71e+01
41	108	3.0	455 9	R57771	Mouse nucleobindin.	3.63e+01
42	107	3.0	561 34	M63043	Streptococcus uberis	4.14e+01
43	106	3.0	708 2	R08402	Am105 encoded by Iida	4.71e+01
44	106	3.0	1325 23	W19540	Mate-enhanced antigen	4.71e+01
45	106	3.0	2482 29	M23996	Human mitotin amino a	4.71e+01

ALIGNMENTS

RESULT 1	
ID	W01044 standard; Protein; 501 AA.
AC	W01044;
DT	28-DEC-1996 (first entry)
DE	Y. pestis FI/V antigen fusion.
KW	Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW	FI antigen; cafI.
OS	Chimeric Yersinia pestis strain GB;
OS	Chimeric synthetic.
FH	Key location/Qualifiers
FT	peptide 1..21
FT	/label= Sig-peptide
FT	/note= "FI antigen signal peptide"
FT	protein 22..170
FT	/label= FI antigen
FT	/note= "mature FI antigen"
FT	peptide 171..176
FT	/label= linker
FT	/note= "6-amino acid peptide linker"
FT	protein 176..501
FT	/label= V antigen
FT	/note= "mature V antigen"
PN	W09628551-A1.
PD	19-SEP-1996.
PR	13-MAR-1996; G00571.
PR	15-SEP-1995; GB-005059.
PR	05-DEC-1995; GB-018946.
PR	05-DEC-1995; GB-024825.
PA	(M1NA) UK SEC FOR DEFENCE.
PI	Bennett AM, Leary SEC, Oyston PCF, Tilball RW, Williamson ED;
DR	WPI; 96-433824/43.
DR	N-PSDB; T38249.
PT	Yersinia pestis V antigen and FI antigen or their protective
PT	epitopic parts - useful in vaccine for protection against plague
PS	Example 3, Page 65-69; 98pp; English.
CC	A fusion protein (W01044) comprises the FI antigen (see also
CC	W01043) and V antigen (see also W01041) of Yersinia pestis joined
CC	by a linker that allows each protein to attain its conformational
CC	state. It is the product of a gene fusion (T38249) obtd. by PCR
CC	amplification of Y. pestis DNA. FI/V fusion protein can be

CC expressed by gut-colonising organism transformants, to induce
CC an immune response against Y. pestis, the causative organism of
CC plague.
CC Sequence 501 AA;

Query Match 92.9%; Score 3294; DB 19; Length 501;
Best Local Similarity 98.4%; Pred. No. 2,076-236;
Matches 493; Conservative 2; Mismatches 3; Indels 3; Gaps 2;

Db 1 mkksivaiatfgtataanaadlactatatltvepariltlykegapitmdngnidt 60
Qy 24 MKKISSVIAIALFGTATANAADLTASTATATLVEPARILTLYEGAPITMDNGNIDT 83
Db 61 ellvgtlllggykgtctstsvnfdaagqmylftsqggnhgftckvigsrdfidis 120
Qy 84 ELLVGTLLLGKYGKGTCTSTSVNFTDAGDPMYLTFTSQDGNHQTFTKVGKSDRFDIS 143
Db 121 pkvgenlvqddvnlatsgqdfvrsigskgklaagkytdavtvtvsnqgsiegritray 180
Qy 144 PKVNEENLVGDVNLATGSQDFVRSIGSKGKLAAGKYTDVTVTVSNQE-F--MIRAY 200
Db 181 eonpqhffiedlekvrveqtlthgssvleelvglvkdnidisikydpkdvsevfannrvt 240
Qy 201 EONPQHFFIEDEKVRVEQTLTGHSVLEELVOLVKDNIDISIKTDPKDVSEVFANRVIT 260
Db 241 ddieilkkilayflpedaillkgyhynqngikrvkefleesspntqwelirafamvmbfs 300
Qy 261 DDIEILKKILAYFLPEDETLTKGSHYDNOQNGIKRVKEFLESSPNTQWELRAFMAVMHFS 320
Db 301 ltadidddilkvlydsnmhbgdarskireelaetaelklysviqaelnkhlssegctin 360
Qy 321 LTADIDDDILKVLVDSNMHGDARSKIREELAEELTAELKLYSVIOAELNKLSSSGTIN 380
Db 361 ihdesinlmdknlvgtytdeeifkasaeykilekmpgtitvgdsekklysvikdfisgenk 420
Qy 381 IHDEKINLMDKNLVGYTDEEIFKASAEYKILEKMPGTITQVGSSEKKLYSINDFLGSENK 440
Db 421 rtgalgnlksynskdhnelshfctcsdksrplndlvsgqtktlqsdltsrfsnaieal 480
Qy 441 RTGALGNLKSNSYNKDNELSHFATTCSDKSRPLNDLVSQTKTQLSDITSRFNSAIEAL 500
Db 481 nrftqkydsvmgrrllldtsqk 501
Qy 501 NRTQKYDSVMORLLDDTSQK 521

RESULT 2
ID W01045 standard; Protein: 480 AA.

AC W01045;
DT 28-DEC-1996 (first entry)
DE Y. pestis Fl/V antigen fusion.
KW Plaque; vaccine; genetic immunisation; V antigen; lcrV;
KW Fl antigen; cagl
OS Chimeric Yersinia pestis strain GB;
OS Chimeric synthetic.
FH Key Location/Qualifiers
FT protein 1..149
FT /label= Fl antigen
FT /note= "mature Fl antigen"
FT peptide 150..155
FT /label= Linker
FT /note= "6-amino acid peptide linker"
FT protein 156..480
FT /label= V antigen
FT /note= "mature V antigen"
PN W09628551-A1.
PD 19-SEP-1996.
PE 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tilball RW, Williamson ED;

DR WPI: 96-433824/43.
DR N-PSDB; T38256.
PR Yersinia pestis V antigen and Fl antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Disclosure: Page 51-55; 98pp; English.

CC A fusion protein (W01045) comprises the Fl antigen (see also
CC W01042) and V antigen (see also W01041) of Yersinia pestis joined
CC by a linker that allows each protein to attain its conformational
CC state. It is the product of a gene fusion (T38256) obtd. by PCR
CC amplification of Y. pestis DNA. Fl/V fusion protein can be
CC expressed by gut-colonising organism transformants, to induce
CC an immune response against Y. pestis, the causative organism of
CC plague.
CC Sequence 480 AA;

Query Match 89.3%; Score 3166; DB 19; Length 480;
Best Local Similarity 98.3%; Pred. No. 1,176-226;
Matches 472; Conservative 3; Mismatches 2; Indels 3; Gaps 2;

Db 1 adlataatatlvepariltlykegapitmdngnidtelvgtlllggykgtctstsv 60
Qy 45 ADLTAATATATLVEPARILTLYKEGAPITMDNGNIDTELVLGTLTGKGTCTSTSV 104
Db 61 nftdaagqmylftsqggnhgftckvigsrdfidispkvgenlvqddvnlatsgq 120
Qy 105 NFTDAAGDPMYLTFTSQDGNHQTFTKVGKSDRFDISPKVNGENLVGDVNLATGSQD 164
Db 121 fvtvrsigskgklaagkytdavtvtvsnqgsiegritayegonpqhffiedlekvrveqtl 180
Qy 155 FVTVRSIGSKGKLAAGKYTDVTVTVSNQE-F--MIRAYEONPQHFFIEDEKVRVEQTLTG 221
Db 181 hgssvleelvglvkdnidisikydpkdvsevfannrvtddieilkkilayflpedaillk 240
Qy 222 HGSSVLEELVOLVKDNIDISIKTDPKDVSEVFANRVITDDIEILKKILAYFLPEDETLTK 281
Db 241 gghydnqilngikrvkefleesspntqwelirafamvmbfsltadidddilkvlydsnmh 300
Qy 282 GGHYDNOQNGIKRVKEFLESSPNTQWELRAFMAVMHFSLTADRIDDDILKVLVDSNMH 341
Db 301 gdarskireelaetaelklysviqaelnkhlssegctinlmdknlvgtytdeeif 360
Qy 342 GDARSKIREELAEELTAELKLYSVIOAELNKLSSSGTINIHDKSNLMDKNLVGYTDEI 401
Db 361 fksaeykilekmpgtitvgdsekklysvikdfisgenkrtgalgnlksynskdhnel 420
Qy 402 FKSAEYKILEKMPGTITQVGSSEKKLYSINDFLGSEKKRTGALGNLKSNSYNKDNEL 461
Db 421 shfctcsdksrplndlvsgqtktlqsdltsrfsnaiealnrfiqkydsvmgrrllldtsqk 480
Qy 462 SHFATTCSDKSRPLNDLVSQTKTQLSDITSRFNSAIEALNRFIQKYDSVMORLLDDTSQK 521

RESULT 3
ID W01040 standard; Protein: 329 AA.

AC W01040;
DT 28-DEC-1996 (first entry)
DE Y. pestis V antigen.
KW Plaque; vaccine; genetic immunisation; V antigen; lcrV;
KW Fl antigen.
OS Yersinia pestis strain GB.
FH Key Location/Qualifiers
FT peptide 1..4
FT /note= "vector-encoded peptide"
PN W09628551-A1.
PD 19-SEP-1996.
PE 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tilball RW, Williamson ED;
DR WPI: 96-433824/43.
DR N-PSDB; T38242.

Seq	Sequence	329 AA:	63.2%:	Score 224.7:	DB 15:	Length 329:
	Query Match		63.2%:	Score 224.7:	DB 15:	Length 329:
	Best Local Similarity	99.4%:	Pred. No. 2.15e-156:			
	Matches 326:	Conservative	1:	Mismatches 0:	Indels 1:	Gaps 1:
Db	3 ef-irayeqnpgpfiedlekvrveqitlghssvyleelvgdvkknidiskypirkdse	61				
QY	194 EFMIKRAEQPHFIEDLEKVRVEQLTGHSVSYLEELVQVKKKNIDISKYPRRDSEV	253				
Db	62 farrvttddellkklilayfpedalikgghydnqigndikrckeflessnptqwejraf	121				
QY	254 FARRVTTDDLELLKILAYFPEDTILIKGCHYNOLONGIKRKVELESSPNTQWEJRAF	313				
Db	122 mawhisltdadridddllkvivdsmbhgdarsklirelaeltaelklysvigaalnkh	181				
QY	374 SSSGGINIHOKSINLMDKNLYGYTDEIFKASAEYKILEMPTTIOVDSEKKIYSIKD	433				
Db	242 flsseqkrtbalnklksvysynkdmelnshfatscdskrpndlyvsqgkrtqdsitrf	301				
QY	434 FLSEENKRTGALGNLKNKSIYNKDNNELSHFATTCSDKSRPLNDLVSQKTTQLSDITSRF	493				
Db	302 nsalealnrfiqkydsvmqgflldatsgk 329					
QY	494 NSALEALNRFIOKRYDSVMQRLDDTSGK 521					
RESULT	5					
ID	W01041: standard; Protein: 329 AA.					
AC	W01041:					
DT	28-DEC-1996 (first entry)					
DE	Y. pestis V antigen.					
KW	Plague; vaccine; genetic immunisation; V antigen; lcrV;					
KW	Fl antigen.					
OS	Yersinia pestis strain GB.					
FT	Key					
FT	peptide					
FT	1..4					
FT	Location/Qualifiers					
FT	/note= "vector-encoded peptide"					
PN	W09628551-A1.					
PD	19-SEP-1996.					
PF	13-MAR-1996: G00571.					
PR	13-MAR-1995: GB-005059.					
PR	15-SEP-1995: GB-018946.					
PR	05-DEC-1995: GB-024825.					
PA	(MINA) UK SEC FOR DEFENCE.					
PI	Bennett AM, Leary SEC, Oyston PCF, Tiltball RW, Williamson ED;					
PI	WPI: 96-433824/43.					
DR	N-PSDB: T38243.					
CC	Yersinia pestis V antigen and Fl antigen or their protective					
PT	epitopic parts - useful in vaccine for protection against plague					
PS	Disclosure: Page 32-35; 98p; English.					
CC	Yersinia pestis V antigen (W01041) is capable of evoking protective					
CC	immune responses in animals. A combined vaccine of V and Fl					
CC	antigens (see also W01042) can at least match the protection					
CC	afforded by live attenuated EV76 vaccine without any of the hazards					
CC	that have kept the EV vaccine from general use. The V antigen is					
CC	produced using the lcrV gene (see also T38243) obtd. from Y. pestis					
CC	by PCR amplification. It can also be prepd. as a fusion with Fl					
CC	antigen (see also W01044-45) and expressed by attenuated Arda or C					
CC	Salmonella typhi and gut-colonising bacteria for protection against					
CC	plague.					
SQ	Sequence 329 AA:					
	Query Match	63.1%:	Score 223.6:	DB 19:	Length 329:	
	Best Local Similarity	99.7%:	Pred. No. 6.15e-156:			
	Matches 324:	Conservative	1:	Mismatches 0:	Indels 0:	Gaps 0:
Db	5 irayeqnpgpfiedlekvrveqitlghssvyleelvgdvkknidiskypirkdsefan	64				

QY 197 IFAYEONPOHFIEDLEKXVNEVDLTGSSVLEELVOLVVKDNIDISIKDKPRKDSVFAN 256
DB 65 rvtlddiellkkllyfjpedailkxghydnqlngiklvkefleesspntqweiafmay 124
QY 257 RVTDDIELLKKILAYFLPEDITILKGYHNOQNGIKRVEKEFLESSPNTQWELRAFMAY 316
DB 125 mhfsitadidddllkvtidsmmhghdarsklreeaeltaeklysvigaetnklhss 184
QY 317 MHFSLTADRIDDDILKVIDSMNHGDARSKLREELAEELTAELKITYSVIOAEINKHLSSS 376
DB 185 gfinhdskeinmdknlvgtdaeifkasaeykilekmpgtiqvqgsekkivsiikflg 244
QY 377 GTINIHDKSINLMDKNLYGTDEIFRKASAEKYLEKMPQTTIIOVDSSEKKIVSINDFLG 436
DB 245 senktgtalgnlksysynkdmeishfctcsksrplndivsgktqlditsrfnsa 304
QY 437 SENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDIVSOKTQLOLSDITSRFNSA 496
DB 305 lealnrfikkydsvmgriiddtsqk 329
QY 497 IEALNRFIOKDYVMORLLDDTSGK 521

RESULT 6
ID R79962 standard; Protein: 329 AA.
AC R79962:
DT 18-APR-1996 (first entry)
DE Partial lcrv (V antigen) of Y. pestis.
KW lcrv; V antigen; virulence; plague; vaccine; epitope.
OS Yersinia pestis.
PN WO924475-A1.
PD 14-SEP-1995.
PF 06-MAR-1995; G00481.
PR 08-MAR-1994; GB-004577.
PA (MINA) UK SEC FOR DEFENCE.
PI Leary SEC, Titball RW, Williamson ED, Leary SE;
DR WPI: 95-328266/42.
DR N-PSDB: T04223.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PS oral or parenteral vaccines for protection against plague
CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrv) encoding
CC all or a protective epitopic part of the mature V protein of Yersinia
CC pestis. The protein was expressed as a fusion protein with maltose
CC binding protein or glutathione-S-transferase in 3 different plasmid
CC vectors. Y. pestis is the highly virulent causative organism of plague
CC in a wide range of animals, including man. The V antigen (lcrv) is an
CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
CC The V antigen is postulated to act as a virulence antigen, and
CC transformed microorganisms cony. recombinant DNA encoding a V antigen
CC protein/peptide are useful in vaccines to protect against plague.
SQ Sequence 329 AA;

Query Match 63.0%; Score 2233; DB 15; Length 329;
Best Local Similarity 99.4%; Pred. No. 1.04e-155;
Matches 323; Conservatve 2; Mismatches 0; Indels 0; Gaps 0;

DB 5 ifayeqnpgfhfiedlekxvnevdltgssvleelvolvdknidisikdkprkdsvefan 64
QY 197 IFAYEONPOHFIEDLEKXVNEVDLTGSSVLEELVOLVVKDNIDISIKDKPRKDSVFAN 256
DB 65 rvtlddiellkkllyfjpedailkxghydnqlngiklvkefleesspntqweiafmay 124
QY 257 RVTDDIELLKKILAYFLPEDITILKGYHNOQNGIKRVEKEFLESSPNTQWELRAFMAY 316
DB 125 mhfsitadidddllkvtidsmmhghdarsklreeaeltaeklysvigaetnklhss 184
QY 317 MHFSLTADRIDDDILKVIDSMNHGDARSKLREELAEELTAELKITYSVIOAEINKHLSSS 376
DB 185 gfinhdskeinmdknlvgtdaeifkasaeykilekmpgtiqvqgsekkivsiikflg 244
QY 377 GTINIHDKSINLMDKNLYGTDEIFRKASAEKYLEKMPQTTIIOVDSSEKKIVSINDFLG 436

DB 245 senktgtalgnlksysynkdmeishfctcsksrplndivsgktqlditsrfnsa 304
QY 437 SENKRTGALGNLKNYSYNKDNNELSHFATTCSDKSRPLNDIVSOKTQLOLSDITSRFNSA 496
DB 305 lealnrfikkydsvmgriiddtsqk 329
QY 497 IEALNRFIOKDYVMORLLDDTSGK 521

RESULT 7
ID W01043 standard; Protein: 170 AA.
AC W01043:
DT 28-DEC-1996 (first entry)
DE Y. pestis F1 antigen (including signal peptide).
KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;
OS Yersinia pestis strain GB.
FH key Location/Qualifiers
FT peptide 1.21
PN WO9628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 13-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 96-433824/43.
DR N-PSDB: T38248.
PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Example 2; Page 61-62; 98pp; English.
CC The F1 antigen (W01043), including the signal peptide, of Yersinia
CC pestis was produced from a DNA sequence (T38248) obtd. by PCR
CC amplification (see also T38257-58) of Y. pestis DNA. Expression
CC of the F1 antigen (see also W01042) by gut-colonising organisms in
CC the form of live vaccines can be used to protect an animal,
CC including humans, against plague.
SQ Sequence 170 AA;

Query Match 31.0%; Score 1100; DB 19; Length 170;
Best Local Similarity 100.0%; Pred. No. 3.31e-70;
Matches 170; Conservatve 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 mkhissvlaiaifgfiatanaadlastactatlveparitltkyegapitmdnuidt 60
QY 24 MKKISSVLAIALFGITATANAADLTASTITATVLEPARITLTLYEGAPITIMDNGNIDT 83
DB 61 ellvgtlilgykktgtsvntfdaagdpmylftsgdgmngfttkvigsdrdfis 120
QY 84 ELLVGTILTLGKYGKGTSTSVNFTDAAGDPMYLFTSGDGMNHQTTTVIGKDSRDFIS 143
DB 121 pkyvgenlyvgddvvlatsqgdfivrsvsgskgklaagkyttdavvtvsnq 170
QY 144 PKVNGENLVGDDVVLATSGQDFIVRSVSGSKGKLAAGRYTDAVVTVSNQ 193

RESULT 8
ID R76528 standard; Protein: 170 AA.
AC R76528:
DT 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen.
KW Vaccine; plague; Salmonella typhimurium; Salmonella typhi;
OS Yersinia pestis.
PN WO9518231-A1.
PD 06-JUL-1995.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA) UK SEC FOR DEFENCE.
PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 95-246396/32.

DR N-PSDB: Q92819.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.
PS Disclosure: Page 20; 27pp; English.
CC The sequence represents the Y. pestis cafi (F1) antigen expressed
CC from plasmid pFORF1b. The DNA construct can be used to transform
CC human or animal gut colonizing microorganisms, specifically
CC attenuated *Salmonella typhimurium* or *Salmonella typhi*. The
CC transformed microorganisms can be used as live/attenuated vaccines
CC which induce immune responses at mucosal surfaces. The vaccines
CC provide protection against infection with Y. pestis, and are
CC parenterally and orally active vaccines offering protection
CC against bubonic and pneumonic plague.
Sequence 170 AA;
Query Match 31.0%; Score 1100; DB 14; Length 170;
Best Local Similarity 100.0%; Pred. No. 3,31e-70;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mkksissvialafgfiatanaadlaastatataleparitltkyegapitmdgnid 60
OY 24 MKKISSVIAIAFGFIATANADLTAATATATLVEPARITLTKEGAPITMDGNIDT 83
Db 61 ellvgtlllggykgttsvntfdaagdmlylftsgdgnhgttkvigsrddfidis 120
OY 84 ELLVGTLLLGKYGKGTSTSVNFTDAAGDPMYLFPTSDGNHGHFTTKVIGKDSRDFIS 143
Db 121 pkvgenlvgdvvlatsgdfivrsgskgklaagkytldavtyvsng 170
OY 144 PKVNGENLVGDVVLATSGDFIVRSIGSKGKLAAGKYTDAVTVTVSNQ 193

RESULT 9
ID W59783 standard; Protein: 170 AA.
AC W59783;
DT 26-OCT-1998 (first entry)
DE Amino acid sequence of F1 antigen nypflsec510.
KM F1 antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
FS Key location/Qualifiers
FT CDS 17..532
FT /*tag= a
FT /product= "F1 antigen"
PN WO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR N-PSDB: V41596.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Disclosure: Pages 53-54; 75pp; English.
CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
Sequence 170 AA;
Query Match 31.0%; Score 1100; DB 34; Length 170;
Best Local Similarity 100.0%; Pred. No. 3,31e-70;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mkksissvialafgfiatanaadlaastatataleparitltkyegapitmdgnid 60
OY 24 MKKISSVIAIAFGFIATANADLTAATATATLVEPARITLTKEGAPITMDGNIDT 83
Db 61 ellvgtlllggykgttsvntfdaagdmlylftsgdgnhgttkvigsrddfidis 120
OY 84 ELLVGTLLLGKYGKGTSTSVNFTDAAGDPMYLFPTSDGNHGHFTTKVIGKDSRDFIS 143

Db 121 pkvgenlvgdvvlatsgdfivrsgskgklaagkytldavtyvsng 170
OY 144 PKVNGENLVGDVVLATSGDFIVRSIGSKGKLAAGKYTDAVTVTVSNQ 193
RESULT 10
ID W59782 standard; Protein: 170 AA.
AC W59782;
DT 26-OCT-1998 (first entry)
DE Amino acid sequence of F1 antigen pypflsec170.
KM F1 antigen; plasmid; vaccine; plague.
OS Yersinia pestis.
FS Key location/Qualifiers
FT CDS 17..532
FT /*tag= a
FT /product= "F1 antigen"
PN WO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR N-PSDB: V41596.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 10; Page 52; 75pp; English.
CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
Sequence 170 AA;
Query Match 31.0%; Score 1100; DB 34; Length 170;
Best Local Similarity 100.0%; Pred. No. 3,31e-70;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mkksissvialafgfiatanaadlaastatataleparitltkyegapitmdgnid 60
OY 24 MKKISSVIAIAFGFIATANADLTAATATATLVEPARITLTKEGAPITMDGNIDT 83
Db 61 ellvgtlllggykgttsvntfdaagdmlylftsgdgnhgttkvigsrddfidis 120
OY 84 ELLVGTLLLGKYGKGTSTSVNFTDAAGDPMYLFPTSDGNHGHFTTKVIGKDSRDFIS 143
Db 121 pkvgenlvgdvvlatsgdfivrsgskgklaagkytldavtyvsng 170
OY 144 PKVNGENLVGDVVLATSGDFIVRSIGSKGKLAAGKYTDAVTVTVSNQ 193
RESULT 11
ID W59788 standard; Protein: 149 AA.
AC W59788;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen pypflmat149.
KM F1 antigen; plasmid; vaccine; plague.
OS Yersinia pestis.
FS Key location/Qualifiers
FT CDS 17..532
FT /*tag= a
FT /product= "F1 antigen"
PN WO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR N-PSDB: V41609.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 10; Page 63; 75pp; English.
CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
Sequence 149 AA;

Query Match 27.4%; Score 971; DB 34; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1,42e-60;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 adltaataatcattvparitltykegapitindgnidtelvgtlilggyktgtstst 60
 OY 45 ADLTAATATATATVLPARITLTYKEGAPITINDGNIDELVGLTILGGYKTGTST 104
 DB 61 nftdaagdpmylfttsqdggnhftkvikgsdrdfdispkvngenlvgddvlatgsqd 120
 OY 105 NFTDAAGDPMYLTFTSQDGNHGFTRKVIKGSDFDISPKVNGENLVGDDVVLATGSQD 164
 DB 121 ffyrslgskgkllaagkytdavrvtsnq 149
 OY 165 FFYRSLGSKGKLLAAGKYTDVAVRVTSNQ 193

RESULT 12
 ID W59787 standard; Protein: 150 AA.
 AC W59787;
 DT 26-OCT-1998 (first entry)
 DE Amino acid sequence of F1 antigen pYpF1mat150.
 KW F1 antigen; plasmid; vaccine; plague.
 OS Yersinia pestis.
 PN MO9824912-A2.
 PD 11-JUN-1998;
 PE 04-DEC-1997; U22617.
 PR 04-DEC-1996; US-767115.
 PA (HESK-) HESKA CORP.
 PI Haanes EJ, Osorio JE, Thomas RE;
 DR MPI: 98-33333/29.
 DR N-PSDB: V41600.
 PT Protection of animals against plague - using nucleic acid encoding
 PT antigen from Yersinia, Pasteurella and Francisella spp.
 PS Claim 10; Page 60; 75pp; English.
 CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
 CC in the method of the invention. Plasmid and host cells are used to
 CC produce recombinant antigens, especially Yersinia pestis antigens.
 CC The recombinant antigens can be used in vaccines that are capable of
 CC protecting an animal from contracting plague.
 SQ Sequence 150 AA;

Query Match 27.4%; Score 971; DB 34; Length 150;
 Best Local Similarity 100.0%; Pred. No. 1,42e-60;
 Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2 adltaataatcattvparitltykegapitindgnidtelvgtlilggyktgtstst 61
 OY 45 ADLTAATATATVLPARITLTYKEGAPITINDGNIDELVGLTILGGYKTGTST 104
 DB 62 nftdaagdpmylfttsqdggnhftkvikgsdrdfdispkvngenlvgddvlatgsqd 121
 OY 105 NFTDAAGDPMYLTFTSQDGNHGFTRKVIKGSDFDISPKVNGENLVGDDVVLATGSQD 164
 DB 122 ffyrslgskgkllaagkytdavrvtsnq 150
 OY 165 FFYRSLGSKGKLLAAGKYTDVAVRVTSNQ 193

RESULT 13
 ID W01042 standard; Protein: 151 AA.
 AC W01042;
 DT 28-DEC-1996 (first entry)
 DE Y. pestis F1 antigen.
 KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;
 KW V antigen
 OS Yersinia pestis strain GB.
 PN W09628531-A1.
 PD 19-SEP-1996;
 PE 13-MAR-1996; G00571.
 PR 13-MAR-1995; GB-005059.
 PR 15-SEP-1995; GB-018946.
 PR 05-DEC-1995; GB-024825.

PA (MINA) UK SEC FOR DEFENCE.
 PI Bennett AM, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
 DR MPI: 96-433824/43.
 DR N-PSDB: T38244.
 PT Yersinia pestis V antigen and F1 antigen or their protective
 PT epitopic parts - useful in vaccine for protection against plague
 PS Disclosure; Page 43-45; 98pp; English.
 CC Yersinia pestis F1 antigen (W01042) is capable of evoking protective
 CC immune responses in animals. A combined vaccine of F1 and V
 CC antigens (see also W01040-41) can at least match the protection
 CC afforded by live attenuated EV76 vaccine without any of the hazards
 CC that have kept the FV vaccine from general use. The F1 antigen is
 CC produced using the cafi gene (see also T38244) obtd. from Y.
 CC pestis by PCR amplification. It can also be prep. as a fusion with
 CC V antigen (see also W01044-45) and expressed by attenuated AroA or
 CC C Salmonella typhi as a live vaccine for long-term protection
 CC against plague. Expression by gut-colonising bacterial transformants
 CC produces a protective response against Y. pestis.
 SQ Sequence 151 AA;

Query Match 27.3%; Score 968; DB 19; Length 151;
 Best Local Similarity 98.0%; Pred. No. 2,38e-60;
 Matches 148; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 1 ssadltaataatcattvparitltykegapitindgnidtelvgtlilggyktgtstst 60
 OY 43 SSADLTAATATATVLPARITLTYKEGAPITINDGNIDELVGLTILGGYKTGTST 102
 DB 61 svnftdaagdpmylfttsqdggnhftkvikgsdrdfdispkvngenlvgddvlatgs 120
 OY 103 SVNFTDAAGDPMYLTFTSQDGNHGFTRKVIKGSDFDISPKVNGENLVGDDVVLATGS 162
 DB 121 qdfvrslyskgkllaagkytdavrvtsnq 151
 OY 163 QDFVRSISGSKGKLLAAGKYTDVAVRVTSNQ 193

RESULT 14
 ID R76526 standard; Protein: 151 AA.
 AC R76526;
 DT 17-DEC-1995 (first entry)
 DE Yersinia pestis cafi (F1) antigen.
 KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
 KW bubonic plague; pneumonic plague.
 OS Yersinia pestis.
 PN W09518231-A1.
 PD 06-JUL-1995;
 PE 23-DEC-1994; G02818.
 PR 24-DEC-1993; GB-026425.
 PA (MINA) UK SEC FOR DEFENCE.
 PI Howells A, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
 DR MPI: 95-246396/32.
 DR N-PSDB: Q92817.
 PT DNA constructs capable of transforming microorganisms - which can be
 PT used as live or attenuated vaccines which induce an immune response.
 PT against Yersinia pestis, at mucosal surfaces.
 PS Disclosure; Page 16; 27pp; English.
 CC The sequence represents the Y. pestis cafi (F1) antigen expressed
 CC from plasmid pFGAL2a. The DNA construct can be used to transform
 CC human or animal gut colonizing microorganisms, specifically
 CC attenuated Salmonella typhimurium or Salmonella typhi. The
 CC transformed microorganisms can be used as live/attenuated vaccines
 CC which induce immune responses at mucosal surfaces. The vaccines
 CC provide protection against infection with Y. pestis, and are
 CC parenterally and orally active vaccines offering protection
 CC against bubonic and pneumonic plague.
 SQ Sequence 151 AA;

Query Match 27.3%; Score 968; DB 14; Length 151;
 Best Local Similarity 98.0%; Pred. No. 2,38e-60;
 Matches 148; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 1 ssadltaataatcattvparitltykegapitindgnidtelvgtlilggyktgtstst 60

WIDEVIEW (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Nov 27 15:37:17 1999; MasPar time 8.24 Seconds

Tabular output not generated. 753.590 Million cell updates/sec

Title: >US-08-699-716A-2
Description: (1-521) from US08699716A.pep
Perfect Score: 3546
Sequence: 1 MGHNNHHHHSSGHIDDD.....RFIOKYSVMQRLDDTSGK 521

Scoring table: PAM 150
Gap 11

Searched: 122461 seqs, 11912985 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCIT9_COMB 4:backfiles1

Statistics: Mean 33.116; Variance 187.062; scale 0.177

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	150	4.2	462	2	US-08-405-Sequence 26, Applicat	3.84e-02
2	150	4.2	462	1	US-08-480-Sequence 26, Applicat	3.84e-02
3	144	4.1	21	2	US-08-651-Sequence 21, Applicat	9.19e-02
4	146	4.1	23	2	US-08-405-Sequence 24, Applicat	6.87e-02
5	146	4.1	23	1	US-08-480-Sequence 24, Applicat	6.87e-02
6	146	4.1	53	2	US-08-651-Sequence 19, Applicat	6.87e-02
7	146	4.1	54	2	US-08-651-Sequence 23, Applicat	6.87e-02
8	146	4.1	323	2	US-08-591-Sequence 5, Applicatio	6.87e-02
9	123	3.5	168	1	US-08-460-Sequence 2, Applicatio	1.82e+00
10	106	3.0	174	3	PCT-US95-0Sequence 2, Applicatio	1.87e+01
11	106	3.0	174	2	US-08-719-Sequence 2, Applicatio	1.87e+01
12	106	3.0	174	1	US-08-261-Sequence 2, Applicatio	1.87e+01
13	106	3.0	2482	1	US-08-328-Sequence 6, Applicatio	1.87e+01
14	105	3.0	3248	3	PCT-US95-1Sequence 1, Applicatio	2.14e+01
15	105	3.0	3248	1	US-08-353-Sequence 1, Applicatio	2.14e+01
16	103	2.9	257	1	US-07-781-Sequence 2, Applicatio	2.79e+01
17	98	2.8	17	2	US-08-600-Sequence 13, Applicat	5.39e+01
18	98	2.8	380	1	US-08-670-Sequence 4, Applicatio	5.39e+01
19	98	2.8	380	1	US-08-585-Sequence 4, Applicatio	5.39e+01
20	98	2.8	380	2	US-08-977-Sequence 4, Applicatio	5.39e+01
21	98	2.8	381	2	US-08-670-Sequence 2, Applicatio	5.39e+01
22	98	2.8	381	2	US-08-977-Sequence 2, Applicatio	5.39e+01
23	98	2.8	381	2	US-08-786-Sequence 3, Applicatio	5.39e+01

24	98	2.8	381	1	US-08-585-Sequence 2, Applicatio	5.39e+01
25	98	2.8	390	2	US-08-786-Sequence 1, Applicatio	5.39e+01
26	100	2.8	793	1	US-08-015-Sequence 3, Applicatio	4.14e+01
27	100	2.8	802	1	US-08-015-Sequence 1, Applicatio	4.14e+01
28	99	2.8	3567	2	US-07-642-Sequence 4, Applicatio	4.73e+01
29	96	2.7	180	1	US-08-328-Sequence 8, Applicatio	6.99e+01
30	96	2.7	183	1	US-08-157-Sequence 6, Applicatio	6.99e+01
31	96	2.7	183	3	PCT-US95-1Sequence 7, Applicatio	6.99e+01
32	96	2.7	183	3	PCT-US95-0Sequence 20, Applicatio	6.99e+01
33	97	2.7	302	2	US-08-203-Sequence 4, Applicatio	6.14e+01
34	97	2.7	302	3	PCT-US95-0Sequence 4, Applicatio	6.14e+01
35	97	2.7	303	3	PCT-US95-0Sequence 2, Applicatio	6.14e+01
36	97	2.7	303	2	US-08-203-Sequence 2, Applicatio	6.14e+01
37	96	2.7	303	2	US-08-961-Sequence 2, Applicatio	6.99e+01
38	97	2.7	414	3	PCT-US92-0Sequence 2, Applicatio	6.14e+01
39	95	2.7	544	2	US-08-694-Sequence 10, Applicati	7.96e+01
40	95	2.7	699	2	US-08-694-Sequence 16, Applicati	7.96e+01
41	97	2.7	763	2	US-08-677-Sequence 2, Applicatio	6.14e+01
42	95	2.7	926	1	US-08-387-Sequence 6, Applicatio	7.96e+01
43	95	2.7	977	2	US-08-694-Sequence 8, Applicatio	7.96e+01
44	95	2.7	1069	1	US-07-777-Sequence 9, Applicatio	7.96e+01
45	96	2.7	1338	2	US-08-728-Sequence 9, Applicatio	6.99e+01

ALIGNMENTS

RESULT	1	STANDARD;	PRT:	462 AA.
ID	US-08-405-496A-26			
XX	xxxxxx			
AC				
DT				
XX				
DE	Sequence 26, Application US/08405496A			
CC	Sequence 26, Application US/08405496A			
CC	Patent No. 5919665			
CC	GENERAL INFORMATION:			
CC	APPLICANT: WILLIAMS, JAMES A.			
CC	TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM			
CC	TITLE OF INVENTION: NEUROTOXIN			
CC	NUMBER OF SEQUENCES: 30			
CC	CORRESPONDENCE ADDRESS:			
CC	ADDRESSEE: MEDLEN & CARROLL, LLP			
CC	STREET: 220 MONTGOMERY STREET, SUITE 2200			
CC	CITY: SAN FRANCISCO			
CC	STATE: CALIFORNIA			
CC	COUNTRY: USA			
CC	ZIP: 94104			
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE: Floppy disk			
CC	OPERATING SYSTEM: PC-DOS/MS-DOS			
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30			
CC	CURRENT APPLICATION DATA:			
CC	APPLICATION NUMBER: US/08/405,496A			
CC	FILING DATE: 16-MAR-1995			
CC	CLASSIFICATION: 424			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 08/329,154			
CC	FILING DATE: 25-OCT-1994			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 08/161,907			
CC	FILING DATE: 02-DEC-1993			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 07/985,321			
CC	FILING DATE: 04-DEC-1992			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 07/429,791			
CC	FILING DATE: 31-OCT-1989			
CC	ATTORNEY/AGENT INFORMATION:			
CC	NAME: INGOLTA, DIANE E.			
CC	REGISTRATION NUMBER: 40,027			

REFERENCE/DOCKET NUMBER: OPHD-01308
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 462 AA: 53932 MW: 1087243 CN;
SQ
Query Match 4.2%; Score 150; DB 2; Length 462;
Best Local Similarity 59.4%; Pred. No. 3,84e-02;
Matches 19; Conservative 7; Mismatches 4; Indels 2; Gaps 1;
Db 1 MGNHHHHHHHSGHIEG--RHMASMARLS 30
QY 1 MGNHHHHHHHSGHIDDDDKMKKISSVIA 32
RESULT 2
ID US-08-480-604A-26 STANDARD; PRT; 462 AA.
XX
AC xxxxxx
DT
XX
DE Sequence 26, Application US/08480604A
XX
CC Sequence 26, Application US/08480604A
CC Patent No. 5736139
CC GENERAL INFORMATION:
CC APPLICANT: KINK, JOHN A.
CC APPLICANT: THALEY, BRUCE S.
CC APPLICANT: PADNEY, NISHA V.
CC APPLICANT: FIRCA, JOSEPH R.
CC APPLICANT: STAFORD, DOUGLAS C.
CC TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
CC TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MEDLEN & CARROLL, LLP
CC STREET: 220 MONTGOMERY STREET, SUITE 2200
CC CITY: SAN FRANCISCO
CC STATE: CALIFORNIA
CC COUNTRY: UNITED STATES OF AMERICA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/480,604A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/422,711
CC FILING DATE: 14-APR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/405,496
CC FILING DATE: 16-MAR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/329,154
CC FILING DATE: 25-OCT-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/161,907
CC FILING DATE: 02-DEC-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/985,321
CC FILING DATE: 04-DEC-1992
CC PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPHD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 462 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 462 AA: 53932 MW: 1087243 CN;
SQ
Query Match 4.2%; Score 150; DB 1; Length 462;
Best Local Similarity 59.4%; Pred. No. 3,84e-02;
Matches 19; Conservative 7; Mismatches 4; Indels 2; Gaps 1;
Db 1 MGNHHHHHHHSGHIEG--RHMASMARLS 30
QY 1 MGNHHHHHHHSGHIDDDDKMKKISSVIA 32
RESULT 3
ID US-08-651-818A-21 STANDARD; PRT; 21 AA.
XX
AC xxxxxx
DT
XX
DE Sequence 21, Application US/08651818A
XX
CC Sequence 21, Application US/08651818A
CC Patent No. 5948889
CC GENERAL INFORMATION:
CC APPLICANT: de Boer, Piet A.J.
CC APPLICANT: Hale, Cynthia A.
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
CC TITLE OF INVENTION: ANTIMICROBIALS
CC NUMBER OF SEQUENCES: 25
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MEDLEN & CARROLL
CC STREET: 220 Montgomery Street, Suite 2200
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/651,818A
CC FILING DATE:
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Carroll, Peter G.
CC REGISTRATION NUMBER: 32,837
CC REFERENCE/DOCKET NUMBER: CASE-02249
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 705-8410
CC TELEFAX: (415) 397-8338
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 21 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: not relevant
CC TOPOLOGY: not relevant
CC MOLECULE TYPE: peptide

SO SEQUENCE 21 AA; 2539 MW; 2077 CN;
Query Match 4.1%; Score 144; DB 2; Length 21;
Best Local Similarity 94.4%; Pred. No. 9,19e-02;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1 MGNHHHHHHHSSGHIE 18
1 MGNHHHHHHHHHSGHID 18
QY 1 MGNHHHHHHHHHSGHID 18
RESULT 4
ID US-08-405-496A-24 STANDARD; PRT: 23 AA.
XX AC xxxxxx
XX DT
XX DE Sequence 24, Application US/08405496A
XX CC Sequence 24, Application US/08405496A
XX CC Patent No. 5919665
XX CC GENERAL INFORMATION:
CC APPLICANT: WILLIAMS, JAMES A.
CC TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
CC TITLE OF INVENTION: NEUROTOXIN
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MEDLEN & CARROLL, LLP
CC STREET: 220 MONTGOMERY STREET, SUITE 2200
CC CITY: SAN FRANCISCO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/405,496A
CC FILING DATE: 16-MAR-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/329,154
CC FILING DATE: 25-OCT-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/161,907
CC FILING DATE: 02-DEC-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/985,321
CC FILING DATE: 04-DEC-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/429,791
CC FILING DATE: 31-OCT-1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: INGOLTA, DIANE E.
CC REGISTRATION NUMBER: 40,027
CC REFERENCE/DOCKET NUMBER: OPHD-01308
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 705-8410
CC TELEFAX: (415) 397-8338
CC INFORMATION FOR SEQ ID NO: 24:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 23 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 23 AA; 2741 MW; 2386 CN;
Query Match 4.1%; Score 146; DB 2; Length 23;
Best Local Similarity 79.2%; Pred. No. 6.87e-02;

Matches 19; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
Db 1 MGNHHHHHHHSSGHIEG--RHM 22
1 MGNHHHHHHHHHSGHIDDDDKHM 24
QY 1 MGNHHHHHHHHHSGHIDDDDKHM 24
RESULT 5
ID US-08-480-604A-24 STANDARD; PRT: 23 AA.
XX AC xxxxxx
XX DT
XX DE Sequence 24, Application US/08480604A
XX CC Sequence 24, Application US/08480604A
XX CC Patent No. 5736139
XX CC GENERAL INFORMATION:
CC APPLICANT: KINK, JOHN A.
CC APPLICANT: THALLEY, BRUCE S.
CC APPLICANT: PADHYE, NISHA V.
CC APPLICANT: FIRCA, JOSEPH R.
CC APPLICANT: STAFFORD, DOUGLAS C.
CC TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
CC TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MEDLEN & CARROLL, LLP
CC STREET: 220 MONTGOMERY STREET, SUITE 2200
CC CITY: SAN FRANCISCO
CC STATE: CALIFORNIA
CC COUNTRY: UNITED STATES OF AMERICA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/480,604A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/422,711
CC FILING DATE: 14-APR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/405,496
CC FILING DATE: 16-MAR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/329,154
CC FILING DATE: 25-OCT-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/161,907
CC FILING DATE: 02-DEC-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/985,321
CC FILING DATE: 04-DEC-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/429,791
CC FILING DATE: 31-OCT-1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: INGOLTA, DIANE E.
CC REGISTRATION NUMBER: 40,027
CC REFERENCE/DOCKET NUMBER: OPHD-01763
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 705-8410
CC TELEFAX: (415) 397-8338
CC INFORMATION FOR SEQ ID NO: 24:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 23 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Swanson & Bratschun, L.L.C.
CC STREET: 8400 East Prentice Avenue, Suite #200
CC City: Englewood
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: diskette, 3.5 inch, 1.40 MB
CC MEDIUM TYPE: storage
CC COMPUTER: IBM COMPATIBLE
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: WORD PERFECT 6.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/591,989
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Barry J. Swanson
CC REGISTRATION NUMBER: 33,215
CC REFERENCE/DOCKET NUMBER: NEX 49
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (303) 793-3333
CC TELEFAX: (303) 793-3433
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 323
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 323 AA; 33944 MW; 442207 CN;
Qy
Query Match 4.1%; Score 146; DB 2; Length 323;
Best Local Similarity 79.2%; Pred. No. 6,87e-02;
Matches 19; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
Db 17 MGNHHHHHHHSSGHIEG--RHM 38
|||
1 MGNHHHHHHHSSGHIDDDCKM 24
Qy
RESULT 9
ID US-08-460-739-5 STANDARD; PRT; 168 AA.
XX
AC xxxxxx
XX
DT
DE
XX
Sequence 5, Application US/08460739
XX
CC Sequence 5, Application US/08460739
CC Patent No. 5698416
CC GENERAL INFORMATION:
CC APPLICANT: Wolf, Marcia K.
CC APPLICANT: Cassels, Frederick J.
CC APPLICANT: Bell, Brian A.
CC TITLE OF INVENTION: Improved Methods for Production of
CC TITLE OF INVENTION: Antigens Under Control of Temperature-Regulated Promoters
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Glenna Hendricks
CC STREET: 9669 A Main Street
CC CITY: Fairfax
CC STATE: VA
CC COUNTRY: US
CC ZIP: 22031
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/460,739

CC FILING DATE: 20-APR-1995
CC CLASSIFICATION: 514
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hendricks, Glenna M.
CC REGISTRATION NUMBER: 32,535
CC REFERENCE/DOCKET NUMBER: WO1F2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (703) 425-4250
CC TELEFAX: (703) 425-2767
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 168 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: Cα3 protein
CC SEQUENCE 168 AA; 17464 MW; 155638 CN;
Qy
Query Match 3.5%; Score 123; DB 1; Length 168;
Best Local Similarity 23.5%; Pred. No. 1.82e+00;
Matches 40; Conservative 48; Mismatches 72; Indels 10; Gaps 10;
Db 1 MLKIKYLLIGSLAMSSYSIAAGPFLTKELALNVISPAALDITWAPQDNLTISNTG-V 59
|||
24 MKKISSV-IALALFGLTA-TANADLTASTATATVLEPAITLYTEGAPITTMNGNI 81
Qy 60 -SNTLVGVLTLSNTSIDTVSIASVSPTSNGT-VFAHETNNSASFAT-ISTDNANI 116
|||
82 DTELVGLTLGKTKGTTS-TSYNFTDAAGDPYLTFTSDGNNHQFTTVICKDSRDF 140
Qy 117 TLD-KMAGNTIV-KTNGSQLPLNPLKFTTEGNEHLVSGNYRANNTIT 164
|||
141 DISPKVNGENLVGDVVLATGSDGFVRSIGSKGK-LAAGKYTDVATVT 189
Qy
RESULT 10
ID PCT-US95-07748A-2 STANDARD; PRT; 174 AA.
XX
AC xxxxxx
XX
DT
DE
XX
Sequence 2, Application PC/TUS9507748A
XX
CC Sequence 2, Application PC/TUS9507748A
CC GENERAL INFORMATION:
CC APPLICANT: The Regents of the University of California
CC TITLE OF INVENTION: CLONED Borrelia burgdorferi VIRULENCE
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 4225 Executive Square, Suite 1400
CC CITY: La Jolla
CC STATE: California
CC COUNTRY: USA
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/07748A
CC FILING DATE: 16-JUN-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HAILE, Ph.D., LISA A.
CC REGISTRATION NUMBER: 38,347

```

CC      REFERENCE/DOCKET NUMBER: 07419/013MO1 (PD3516)
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (619) 678-5070
CC      TELEFAX: (619) 678-5099
CC      INFORMATION FOR SEQ ID NO: 2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 174 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 174 AA; 20290 MW; 153533 CN;
Cc
Cc      Query Match          3.0%; Score 106; DB 3; Length 174;
Cc      Best Local Similarity 34.5%; Pred. No.1.87e+01;
Cc      Matches 20; Conservative 12; Mismatches 24; Indels 2; Gaps 2
Cc
Db      115 IKGPNISFDYLIQDSDK-IDYAEKYG-E-KARENFEESYKKDKITAVKQILQIQLD 170
Cc      : |:::|::| | | | | | | | | | | | | | | | | | | | | | | | | |
Cc      219 LTGGSSVLEELVQLVYDKNDISIKYDPRKDSVFANRVITDDIELKKILAEFLPE 276
Cc
Cc      RESULT 11
Cc      ID US-08-719-124-2 STANDARD: PRT; 174 AA.
Cc      XX
Cc      XX xxxxxx
Cc
Cc      Sequence 2, Application US/08719124
Cc      CC
Cc      Patent No. 5834395
Cc      CC
Cc      GENERAL INFORMATION:
Cc      CC
Cc      APPLICANT: Champion, Cheryl I.
Cc      CC
Cc      APPLICANT: Lovett, Michael A.
Cc      CC
Cc      APPLICANT: Haake, David A.
Cc      CC
Cc      APPLICANT: Miller, James N.
Cc      CC
Cc      APPLICANT: Bianco, David R.
Cc      CC
Cc      TITLE OF INVENTION: CLONED Borrelia burgdorferi VIRULENCE
Cc      CC
Cc      TITLE OF INVENTION: PROTEIN
Cc      CC
Cc      NUMBER OF SEQUENCES: 8
Cc      CC
Cc      CORRESPONDENCE ADDRESS:
Cc      CC
Cc      ADDRESSEE: Spensley Horn Jubas & Lubitz
Cc      CC
Cc      STREET: 1880 Century Park East, Suite 500
Cc      CC
Cc      CITY: Los Angeles
Cc      CC
Cc      STATE: California
Cc      CC
Cc      COUNTRY: USA
Cc      CC
Cc      ZIP: 90067
Cc      CC
Cc      COMPUTER READABLE FORM:
Cc      CC
Cc      MEDIUM TYPE: floppy disk
Cc      CC
Cc      COMPUTER: IBM PC compatible
Cc      CC
Cc      OPERATING SYSTEM: PC-DOS/MS-DOS
Cc      CC
Cc      SOFTWARE: PatentIn Release #1.0, Version #1.25
Cc      CC
Cc      CURRENT APPLICATION DATA:
Cc      CC
Cc      APPLICATION NUMBER: US/08/719,124
Cc      CC
Cc      FILING DATE: 24-SEP-1996
Cc      CC
Cc      CLASSIFICATION: 530
Cc      CC
Cc      PRIOR APPLICATION DATA:
Cc      CC
Cc      APPLICATION NUMBER: 08/261,825
Cc      CC
Cc      FILING DATE: 17-JUN-1994
Cc      CC
Cc      ATTORNEY/AGENT INFORMATION:
Cc      CC
Cc      NAME: TUMARKIN, LISA A.
Cc      CC
Cc      REGISTRATION NUMBER: P-38,347
Cc      CC
Cc      REFERENCE/DOCKET NUMBER: PD3516
Cc      CC
Cc      TELECOMMUNICATION INFORMATION:
Cc      CC
Cc      TELEPHONE: (619) 455-5100
Cc      CC
Cc      TELEFAX: (619) 455-5110
Cc      CC
Cc      INFORMATION FOR SEQ ID NO: 2:
Cc      CC
Cc      SEQUENCE CHARACTERISTICS:
Cc      CC
Cc      LENGTH: 174 amino acids
Cc      CC
Cc      TYPE: amino acid
Cc      CC
Cc      *TOPOLOGY: linear
Cc      CC
Cc      MOLECULE TYPE: protein
Cc

```

[illegible]

XX		xxxxxx	
XX			
DT			
DE		Sequence 6, Application US/08328254	
CC			
CC		Sequence 6, Application US/08328254	
CC		Patent No. 5710022	
CC		GENERAL INFORMATION:	
CC		APPLICANT: Zhu, Xueliang	
CC		APPLICANT: Lee, Wen-Hwa	
CC		TITLE OF INVENTION: A No. 5710022ei Nuclear Mitotic Phosphoprotein	
CC		NUMBER OF SEQUENCES: 8	
CC		CORRESPONDENCE ADDRESS:	
CC		ADDRESSEE: Campbell and Flores	
CC		STREET: 4370 La Jolla Village Drive, Suite 700	
CC		City: San Diego	
CC		STATE: California	
CC		COUNTRY: USA	
CC		ZIP: 92122	
CC		COMPUTER READABLE FORM:	
CC		MEDIUM TYPE: Floppy disk	
CC		COMPUTER: IBM PC compatible	
CC		OPERATING SYSTEM: PC-DOS/MS-DOS	
CC		SOFTWARE: Patentin Release #1.0, Version #1.25	
CC		CURRENT APPLICATION DATA:	
CC		APPLICATION NUMBER: US/08/328,254	
CC		FILING DATE: 24-OCT-1994	
CC		CLASSIFICATION: 435	
CC		PRIOR APPLICATION DATA:	
CC		APPLICATION NUMBER: US 08/141,239	
CC		FILING DATE: 22-OCT-1993	
CC		ATTORNEY/AGENT INFORMATION:	
CC		NAME: Campbell, Cathryn A.	
CC		REGISTRATION NUMBER: 31,815	
CC		REFERENCE/DOCKET NUMBER: P-CJ 1191	
CC		TELECOMMUNICATION INFORMATION:	
CC		TELEPHONE: (619) 535-9001	
CC		TELEFAX: (619) 535-8949	
CC		INFORMATION FOR SEQ ID NO: 6:	
CC		SEQUENCE CHARACTERISTICS:	
CC		LENGTH: 2482 amino acids	
CC		TYPE: amino acid	
CC		TOPOLOGY: linear	
CC		MOLECULE TYPE: protein	
SO		SEQUENCE 2482 AA; 283493 MW; 30208241 CN;	
		Query Match 3.0%; Score 106; DB 1; Length 2482;	
		Best Local Similarity 18.5%; Pred. No. 1.87e+01;	
		Matches 74; Conservative 119; Mismatches 184; Indels 24; Gaps 22	
Dd	1498	ADEKQLHIAELKRE-REND-SLKDVENL-EKELQISENOEVLILDAENSKAEVYT	1554
Oy	120	SODGNNHOFTTKVGKDSRDPDISRKVNGENLVGDVVATGESODFPVRSGISGGKLAA	179
Dd	1555	LKTOIEEARSLKYVELDVTLRSEKNLTROIOE-KOGOLSELDK-LTSPFKSLEEKE	1612
Oy	180	GKY-IDANTYIVSNPEFMIRAYEQHPQHIEDLEKVRVBOULTGHSSVYLELYOLVKDKN	238
Dd	1613	QAEOIQKEESTYAVEMLONLKELNEAVALACGDEIMKATEQSIDPPIEEHQLRNSTE	1672
Oy	239	-IDISIKYDPDRDSEVFANRV--ITDDIELLKKILALYFLPEDTLKGK-HYDNONGIJK	294
Dd	1673	KLRALLADEKKQCLVDLOLKESEHNADLLKRVLENLEFELEIATNDPHALLEANSKG	1732
Oy	295	RKVEPLESSPTNQW-ELRAFAVAVMHFSLTDADRIDD-DI-LKVIVDSNMHHGDARSKLE	350
Dd	1733	EVELTKA--KTIGMTQSRGLDELVDVTTIRSEKENLTNELOKOEERTISELEINSSFE-NI	1789
Oy	351	ELAEILTALKETISYIOALEINKHLSSGTIIINHDKSINLMKDRLVGYTDBETIRKASNYKI	410
Dd	1790	LOEKEOEVYOMKEKSTAMLEMLOTQLKELNERVALAHNDQACK-AKEON-ISS-QVECL	1846

DB	1847	ELERKAOLLOGLDEAKNNYIV-LQSSVNGLIQVEDGKQKLE	1886
QY	469	S-DKSRPLNDLVSKTQTQSDITSRFSNAIEALNRFTQKYD	508
RESULT	14		
ID	PCT-US95-16216-1	STANDARD;	PRT; 3248 AA.
AC	xxxxxx		
DT			
XX			
XX			
XX			
DE	Sequence 1, Application PC/TUS9516216		
CC	Sequence 1, Application PC/TUS9516216		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Yen, Timothy J.		
CC	APPLICANT: Rattner, Jerome B.		
CC	TITLE OF INVENTION: Nucleic Acid Encoding a Transiently		
CC	TITLE OF INVENTION: Expressed Kinetochores Protein, and Methods of Use		
CC	NUMBER OF SEQUENCES: 4		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: Dann, Dorfman, Herrell and Skillman		
CC	STREET: 1601 Market Street Suite 720		
CC	CITY: Philadelphia		
CC	STATE: PA		
CC	COUNTRY: USA		
CC	ZIP: 19103-2307		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: Patent In Release #1.0, Version #1.30		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: PCT/US95/16216		
CC	FILING DATE:		
CC	CLASSIFICATION:		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 08/353,700		
CC	FILING DATE: 09-DEC-1995		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Reed, Janet E.		
CC	REGISTRATION NUMBER: 36,252		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: (215) 563-4100		
CC	TELEFAX: (215) 563-4044		
CC	INFORMATION FOR SEQ ID NO: 1:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 3248 amino acids		
CC	TYPE: amino acid		
CC	STRANDEDNESS: not relevant		
CC	TOPOLOGY: not relevant		
CC	MOLECULE TYPE: protein		
CC	HYPOTHETICAL: NO		
CC	ANTI-SENSE: NO		
CC	SEQUENCE 3248 AA: 372207 MW: 51689535 CN:		
DB	2226	ADEKKOHLIAEKLKERE-REND-SLKDKVENL-ERELQMSSENOELVILDAENSKAEVET	2282
QY	120	SQDNNHOFITTKVIGKSDRSPDKVNGENLVGDVVLATGSDOFFVRSIGSGKGLAA	179
DB	2283	LKTOIEBARSILKPELDIVLIRSEKELTQIOIE-KOGQISELDK-LLSFSKILBEKE	2340
QY	180	GKY-TDAVTVTSNOEFIRAYEONPOPFIDLEKVRVEQLTGHGSSVLEELVOLVADKN	238
DB	2341	QAEIOIKSESTAVEMLONLQKELNEAVAAACGDOEIKKATQESLDPIIEEHOLRNSIE	2400

```

OY 229 -IDISIKIDPRKDESEVFANRV--ITDDIELLKKIILAYFLPEDTILKGG-HYDNOLONGIK 294
Db 2401 KLRARLEADEKKOLCVLOOLKESEHHADLLKGRVENLERELEIARTNOEHAALEAENSGK 2460
OY 295 RKKEFLLESSPNTQW-ELRAFMAVHFS-LTADRIDD-DI-LKVIYDSMNHGDARSKLRE 350
Db 2461 EVELTKA--KIEGMOQSLRGLELDVVTIRSEKENLTNELOKEOERISELEIINSFE-NI 2517
OY 351 ELAELTLAELKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGTYDDEIFKASAEYKI 410
Db 2518 LOEKOEKVOOMEKKSSTAMEMLOTOLKELNEVAALHNDQECK-AKEON-LSS-QVECL 2574
OY 411 LEKMPOTTIOV-DGSEKKIVSIKDFLGSEKNRTGALGNLKNKSYNKNDELSPHATTC- 468
Db 2575 ELEKAQLOGLDEAK 2589
OY 469 S-DKSRPLNDLVSOEK 482
```

```

RESULT 15
ID US-08-353-700-1 STANDARD; PRT; 3248 AA.
AC xxxxxx
DT
XX
XX
DE Sequence 1, Application US/08353700
XX
CC Sequence 1, Application US/08353700
CC Patent No. 5599919
CC GENERAL INFORMATION:
CC APPLICANT: YEN, TIMOTHY J.
CC APPLICANT: RATNER, JEROME B.
CC TITLE OF INVENTION: NUCLEIC ACID ENCODING A
CC TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
CC TITLE OF INVENTION: AND METHODS OF USE
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: DANN, DOREMAN, HERRELL AND SKILLMAN
CC STREET: 1601 MARKET STREET, SUITE 720
CC CITY: PHILADELPHIA
CC STATE: PA
CC COUNTRY: USA
CC ZIP: 19103-2307
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/353,700
CC FILING DATE: 09-DEC-1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: REED, JANET E.
CC REGISTRATION NUMBER: 36,252
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (215) 563-4100
CC TELEFAX: (215) 563-4044
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 3248 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC HYPOTHETICAL: NO
CC ANTI-SENSE: NO
CC ORIGINAL SOURCE:
CC ORGANISM: HUMAN
CC SEQUENCE 3248 AA; 372207 MW; 51689535 CN;
```

```

Query Match 3.0%; Score 105; DB 1; Length 3248;
Best Local Similarity 18.4%; Pred. No. 2.14e+01;
Matches 69; Conservative 112; Mismatches 171; Indels 23; Gaps 21;
Db 2226 ADEKQOLHAIEKURE-REND-SLKDKVENL-ERELQMSSENOELVTLDAENSKAEVET 2282
OY 120 SODGNNHOFTTIVITGKDSRDPDISPKVNGENLVGDVVLATGSODFFVRSIGSGKGLAA 179
Db 2283 LKTOIEEMARSLKIPELDVLTRSEKENLTQIOE-KOGQISELDK-DLSSEKSLLEEKE 2340
OY 180 GKV-TDAVTVFVSNQEFMIRAVEQNPQHFIEDLEKVRVEQLTGHSSVLEELVQVKN 238
Db 2341 QAEIOKEESKTAVALMDONOLKELNEVAALCGDOEIKKATEQSDPPIEEHOLRNSIE 2400
OY 239 -IDISIKYDPRKDESEVFANRV--ITDDIELLKKIILAYFLPEDTILKGG-HYDNOLONGIK 294
Db 2401 KLRARLEADEKKOLCVLOOLKESEHHADLLKGRVENLERELEIARTNOEHAALEAENSGK 2460
OY 295 RKKEFLLESSPNTQW-ELRAFMAVHFS-LTADRIDD-DI-LKVIYDSMNHGDARSKLRE 350
Db 2461 EVELTKA--KIEGMOQSLRGLELDVVTIRSEKENLTNELOKEOERISELEIINSFE-NI 2517
OY 351 ELAELTLAELKIYSVIOAEINKHLSSSGTINIHDKSINLMDKNLYGTYDDEIFKASAEYKI 410
Db 2518 LOEKOEKVOOMEKKSSTAMEMLOTOLKELNEVAALHNDQECK-AKEON-LSS-QVECL 2574
OY 411 LEKMPOTTIOV-DGSEKKIVSIKDFLGSEKNRTGALGNLKNKSYNKNDELSPHATTC- 468
Db 2575 ELEKAQLOGLDEAK 2589
OY 469 S-DKSRPLNDLVSOEK 482
```

Search completed: Sat Nov 27 15:37:44 1999
Job time : 27 secs.

[illegible]


```
Db 673 accaccattagtgatgagcgagagaaaaaatagcttcgataaagacttcttga 732
OY 1249 ACCACCTTACGTTGAGTGGAGCGAGAAAAAATAGCTTCGATTAAGAGACTTCTTGA 1308
Db 733 agtgaagataaagaacggcgcttggtgaatcgtgaataactcatcttataataa 732
OY 1309 AGTGAGATATAAAAGACCGGGCGCTTGGTATCTGAAAAATCATATCTTTATATATA 1368
Db 793 gataataatgaatatacttcttccacacactgtcggataaagtcaggcgctcaac 852
OY 1369 GATTAATATGATATATCTCACTTTGCCACACCTGCTGGATTAAGTCAGGCGCTCAAC 1428
Db 853 gaacttggtgccaaaaaacaactcaagctgtctgatatatacattcaagtttaactga 912
OY 1429 GACTTGGTTAGCCAAAAAACACCTCCTGATATTCATCACGCTTTTAATTACAGCT 1488
Db 913 attgaagcactgaacgcttcattcagaatatgattcattgtaagcttcgtctagat 972
OY 1489 ATTGAAGCCTGAAACCTTTTCATTCAAGAAATATGATTCACTGATGCAACGCTGCTAGAT 1548
Db 973 gacacgctgtgtaaatga 990
OY 1549 GACACGCTCTGTGAATGA 1566

RESULT 3
ID 092819 standard; DNA; 542 BP.
AC 092819;
DE Yersinia pestis cafI (F1) antigen in plasmid pFOR1b.
KW vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
OS Yersinia pestis.
FH key Location/Qualifiers
FT CDS 2..7
FT /tag= a
FT /note= "first protein encoded by pFOR1b"
FT misc-feature 1..6
FT /tag= b
FT misc-feature 536..541
FT /tag= c
FT /note= "cafI open reading frame downstream seq."
FT CDS 21..530
FT /tag= d
FT /note= "cafI fusion protein"
FT PN M09518231-A1.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PI (MINA ) UK SEC FOR DEFENCE.
PI Howells A, leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 95-246396/32.
DR P-PSDB: R76528.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.
PS Disclosure: Page 19-20; 27pp; English.
CC The sequence represents the plasmid pFOR1b including the entire
CC Y. pestis cafI (F1) antigen gene having a 5' tail including a SacI
CC restriction site, and up to TATAG downstream of the cafI ORF.
CC The DNA construct can be used to transform human or animal gut
CC colonizing microorganisms, specifically attenuated Salmonella
CC typhimurium or Salmonella typhi. The transformed microorganisms
CC can be used as live/attenuated vaccines which induce immune
CC responses at mucosal surfaces. The vaccines provide protection
CC against infection with Y. pestis, and are parenterally and orally
CC active vaccines offering protection against bubonic and pneumonic
CC plague.
SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;
```

Query Match 33.0%; Score 516; DB 15; Length 542;
Best Local Similarity 99.6%; Pred. No. 0.00e+00;

```
Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 19 atatgaaaaaalcagtlccgltatcgcattgcatlatttgaaactatgcaactgca 78
OY 68 ATATGAAAAAAATCACTTCGTTATCGCATTCATTTATTGGAACTATGCAACTGCTA 127
Db 79 atgcgcagatttaactgcaagaccactgcaagcgcaactctgttgaaccccgca 138
OY 128 ATGCGGACAGATTTTAACTGCAAGCACCATGCAACGGCAACTCTTTGTAACGCGCCGA 187
Db 139 tcaacttatacataaggaagcgctccaattacaattatgcaatggaatcgata 198
OY 188 TCACCTTACATATTAAGAGAGCGCTCCAAATTAATGACAAATGGAACATGATGATA 247
Db 199 cagaattactgtgtgtaagccttactctgtgcgcataaaaaagaagacacatagacat 258
OY 248 CAGAAATTACTTGTGTAGCGCTTACTTGGCGGCTATAAAGACGACACATGACAT 307
Db 259 ctgttaactttacagatgcgcgggtgatcccatgtaacttaacttactctagatg 318
OY 308 CTGTTAACTTTACAGATGCGCGCGGTGATCCCATGTAATTAACATTTACTTCAGGATG 367
Db 319 gaaataacacacacactcactacacaaagtattgcaagattcctaagatttgatctc 378
OY 368 GAAATTAACACCAATTCACACAAAAGTATGCGAAGATTTTAAAGATTTGATATCT 427
Db 379 ctccaaagtaaacgggtgaaacctgtgtgggatgacgtctgtctgtcagcgagcc 438
OY 428 CTCCTAAGGTAAACGGGTGAACACTTGTGGGGATACGCTGCTTGGCTACGGGACGC 487
Db 439 aggaattcttctgtcgtcaattgtgtccaaagcgtaaaccttgcaagagtaataca 498
OY 488 AGGATTTCTTTGTTCCCTCATTTGTTCCAAAGCGGTAAACTTCGACAGATTAATACA 547
Db 499 ctgatctgtaacctgaacctgacatctcaacaataacatc 538
OY 548 CTGATGCTGTAAACCGTAACCGTATCTAACCAAGAAATTCAT 587

RESULT 4
ID 092817 standard; DNA; 541 BP.
AC 092817;
DE Yersinia pestis cafI (F1) antigen in plasmid pFGAL2A.
KW vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
OS Yersinia pestis.
FH key Location/Qualifiers
FT CDS 2..454
FT /tag= a
FT misc-feature 1..6
FT /tag= b
FT /note= "lacZ promoter fusion site"
FT misc-feature 536..541
FT /tag= c
FT /note= "vector pFGAL2A bases"
FT PN M09518231-A1.
PF 06-JUL-1995.
PR 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PI (MINA ) UK SEC FOR DEFENCE.
PI Howells A, leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 95-246396/32.
DR P-PSDB: R76526.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.
PS Claim 7; Page 15-16; 27pp; English.
CC The sequence represents the plasmid pFGAL2A construct showing the
CC fusion of the first few bases of beta-galactosidase in the vector
CC with the Y. pestis cafI (F1) antigen minus its signal sequence and
CC having a 5' tail including a SacI restriction site, and up to the
CC cafI AAC-3' end with some vector bases. The DNA construct can be
```


THIS PAGE BLANK (USPTO)